BAB II
TINJAUAN PUSTAKA

2.1 Kajian Pustaka Relevan

Tinjauan penelitian sebelumnya dilakukan untuk memperkuat landasan penelitian dan menentukan posisi penelitian ini. Tinjauan ini memberikan gambaran umum tentang fokus penelitian, teknik yang digunakan, dan hasil utama. Penelitian ini berfokus pada analisis sentimen bahasa Indonesia dengan metode *IndoBERT* dan penggunaan metode SHAP (*SHapley Additive exPlanations*) sebagai pendekatan *Explainable Artificial Intelligence* (XAI) untuk meningkatkan transparansi hasil prediksi model. Penulis dapat menentukan evolusi metode yang digunakan, manfaat dan kekurangan metode, dan posisi penelitian ini sebagai pengembangan dalam bidang analisis sentimen yang dapat diinterpretasikan dengan model *transformer* dengan meninjau penelitian sebelumnya.

Penelitian pertama yang dilakukan oleh (Supriyadi & Makatita, 2025). berjudul *"Sentiment Analysis of TikTok User Comments on QRIS Adoption in Indonesia Using IndoBERT"* menganalisis sentimen pengguna TikTok terhadap adopsi *Quick Response Code Indonesian Standard* (QRIS). Penelitian tersebut menggunakan model *IndoBERT* untuk mengklasifikasikan 1.128 komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif, dengan menerapkan tahapan *preprocessing* yang relatif sederhana. Hasil penelitian menunjukkan bahwa *IndoBERT* memberikan performa klasifikasi yang baik, ditunjukkan oleh nilai *precision*, *recall*, dan *F1-score* yang tinggi serta kemampuan yang stabil dalam menangani variasi bahasa pada media sosial. Distribusi sentimen yang diperoleh terdiri atas 50,9% sentimen positif, 17,9% sentimen netral, dan 31,2% sentimen negatif. Selain itu, penelitian tersebut mengidentifikasi bahwa kemudahan penggunaan dan penerimaan internasional menjadi faktor yang mendorong sentimen positif, sedangkan kendala teknis dan isu keamanan menjadi faktor utama yang memengaruhi sentimen negatif.

Penggunaan model *IndoBERT* untuk melakukan analisis sentimen pada data media sosial berbahasa Indonesia dan penggunaan klasifikasi multi-kelas (positif, negatif, dan netral) adalah hal yang membuat kedua penelitian ini memiliki persamaan. Sumber data yang digunakan berbeda, penelitian (Supriyadi & Makatita, 2025) menggunakan

komentar TikTok yang berkaitan dengan QRIS, sedangkan penelitian ini menggunakan komentar Instagram pada akun resmi Dewan Perdamaian. Selain itu, penelitian ini tidak menggunakan *Explainable Artificial Intelligence* (XAI), khususnya metode SHAP. Namun, metode ini digunakan untuk meningkatkan keterbacaan dan transparansi hasil prediksi model.

Penelitian kedua yang dilakukan oleh (Setiawan et al., 2025) dengan judul *"Implementation of IndoBERT for Sentiment Analysis of the Constitutional Court's Decision Regarding the Minimum Age of Vice Presidential Candidates"* bertujuan untuk memeriksa seberapa baik model *IndoBERT* mengklasifikasikan sentimen komentar YouTube berbahasa Indonesia yang berkaitan dengan keputusan Mahkamah Konstitusi tentang batas usia calon wakil presiden pada tahun 2024. Penelitian ini menggunakan dataset sebanyak 23.796 komentar yang dikumpulkan melalui YouTube Data API. Langkah-langkah prapemrosesan seperti normalisasi, *case folding*, pembersihan teks, penghapusan simbol, penghapusan *stopword*, dan *stemming* digunakan untuk memproses data secara menyeluruh. Untuk label sentimen, ada tiga kategori yaitu positif, negatif, dan netral.

Studi ini membandingkan beberapa model *IndoBERT, BERT, Support Vector Machine* (SVM), dan *Random Forest* dengan pembagian data 80:20 antara data pelatihan dan data uji. Hasilnya menunjukkan bahwa *IndoBERT* memiliki akurasi tertinggi sebesar 95%, mengungguli *BERT* sebesar 92%, SVM sebesar 88%, dan *Random Forest* sebesar 85%. Ini menunjukkan kemampuan *IndoBERT* untuk memahami konteks bahasa Indonesia dengan lebih baik, terutama saat menganalisis sentimen yang kompleks dalam teks media sosial.

Kesamaan penelitian tersebut dengan penelitian ini terletak pada penggunaan model *IndoBERT* untuk analisis sentimen pada data media sosial berbahasa Indonesia, serta penggunaan klasifikasi multi-kelas (positif, negatif, dan netral). Selain itu, kedua penelitian juga berfokus pada isu kebijakan publik yang memicu respons masyarakat di media sosial. Adapun perbedaannya terletak pada platform dan sumber data yang digunakan, di mana penelitian tersebut menggunakan komentar YouTube terkait isu politik, sedangkan penelitian ini menggunakan komentar Instagram pada akun resmi pemerintah terkait kebijakan *Board of Peace*. Selain itu, penelitian tersebut belum

mengintegrasikan pendekatan *Explainable Artificial Intelligence* (XAI) seperti SHAP, yang dalam penelitian ini digunakan untuk memberikan penjelasan terhadap hasil prediksi model sehingga meningkatkan transparansi dan kepercayaan terhadap sistem.

Penelitian ketiga yang dilakukan oleh (Lundberg & Lee, 2017) dengan judul *"A Unified Approach to Interpreting Model Predictions"* mengusulkan suatu kerangka terpadu untuk menginterpretasikan hasil prediksi model *machine learning* yang kompleks. Penelitian ini memperkenalkan metode SHAP (*SHapley Additive exPlanations*), yaitu pendekatan berbasis teori permainan yang memberikan nilai kontribusi pada setiap fitur terhadap suatu prediksi tertentu. SHAP dirancang untuk mengatasi permasalahan interpretabilitas pada model *black box* seperti *ensemble learning* dan *deep learning*, yang umumnya memiliki akurasi tinggi tetapi sulit dijelaskan.

Dalam penelitian tersebut, SHAP mengidentifikasi kelas baru dari metode *additive feature importance* dan membuktikan secara teoritis bahwa terdapat solusi unik dengan sifat-sifat yang diinginkan, seperti konsistensi dan keadilan dalam atribusi kontribusi fitur. Selain itu, SHAP juga mampu menyatukan berbagai metode interpretasi sebelumnya ke dalam satu kerangka yang lebih terstruktur, serta menawarkan peningkatan dari segi kinerja komputasi dan kesesuaian dengan intuisi manusia dalam memahami hasil model.

Kesamaan penelitian tersebut dengan penelitian ini terletak pada penggunaan pendekatan *Explainable Artificial Intelligence* (XAI) untuk meningkatkan interpretabilitas model *machine learning*, khususnya dalam memahami kontribusi fitur terhadap hasil prediksi. Adapun perbedaannya terletak pada fokus penelitian, di mana penelitian tersebut bersifat konseptual dan umum untuk berbagai jenis model, sedangkan penelitian ini mengimplementasikan metode SHAP secara spesifik pada model *IndoBERT* untuk analisis sentimen komentar Instagram terkait kebijakan *Board of Peace*. Dengan demikian, penelitian ini mengadopsi konsep SHAP sebagai dasar dalam menjelaskan hasil prediksi model secara lebih transparan dan aplikatif pada kasus analisis opini publik.

Penelitian keempat yang dilakukan oleh (Fikry Destryanto et al., 2025) dengan judul *"Sentiment Analysis of Public Response to the Free Nutritious Meal Program on Instagram Using IndoBERT"* Tujuan dari studi ini adalah untuk menganalisis tanggapan

publik terhadap kebijakan pemerintah yang dikenal sebagai Program Makanan Bergizi Gratis (MBG) melalui komentar yang dibuat oleh pengguna Instagram. Metode utama untuk klasifikasi sentimen dalam penelitian ini adalah model *IndoBERT*, dan tahapan penelitian mencakup pengumpulan data, prapemrosesan teks, pelabelan sentimen, dan pelatihan dan evaluasi model.

Data yang digunakan dalam penelitian ini berasal dari komentar pengguna pada postingan Instagram yang terkait dengan program MBG. Kumpulan data ini dikurangi dari 8.447 komentar mentah menjadi 3.136 titik data yang valid dan diberi anotasi selama tahap prapemrosesan. Hasil analisis menunjukkan bahwa 40% komentar adalah positif, 33,3% netral, dan 26,7% negatif. Hal ini menunjukkan bahwa mayoritas masyarakat cenderung bersikap netral terhadap program tersebut, meskipun beberapa pengguna masih mengkritiknya.

Dengan tingkat akurasi 99,89% pada data pelatihan dan 93,95% pada data uji, model *IndoBERT* menunjukkan kinerja yang sangat baik. Selain itu, analisis menggunakan matriks kebingungan menunjukkan bahwa kesalahan klasifikasi relatif kecil dan umumnya terjadi pada kelas sentimen netral dan negatif, nilai presisi, recall, dan skor F1 juga menunjukkan konsistensi yang tinggi dalam klasifikasi sentimen. Hasilnya menunjukkan bahwa *IndoBERT* berfungsi dengan baik untuk menganalisis sentimen pada teks berbahasa Indonesia, terutama data media sosial informal.

Penggunaan model *IndoBERT* dan klasifikasi multi-kelas (positif, negatif, dan netral) untuk melakukan analisis sentimen terhadap komentar Instagram yang berkaitan dengan kebijakan pemerintah menunjukkan kesamaan antara penelitian tersebut dan penelitian ini. Penelitian ini berfokus pada kebijakan Dewan Perdamaian, sedangkan penelitian sebelumnya berfokus pada Program Makanan Bergizi Gratis (MBG). Selain itu, penelitian ini tidak menggunakan pendekatan Kecerdasan Buatan yang Dapat Dijelaskan (XAI) seperti SHAP, pendekatan ini digunakan untuk memberikan penjelasan tentang prediksi model, meningkatkan transparansi dan kepercayaan terhadap hasil analisis.

Penelitian kelima yang dilakukan oleh (Athfaal et al., 2025) dengan judul *"Public Sentiment Analysis of Marketplace Tax Policy Using IndoBERT"* bertujuan untuk menganalisis sentimen publik terhadap kebijakan pajak marketplace melalui data komentar pengguna Instagram. Penelitian ini menggunakan model *IndoBERT* dengan

dataset sebanyak 1.150 komentar yang dikumpulkan melalui teknik *web scraping*. Proses penelitian dilakukan dengan pendekatan SEMMA (*Sample, Explore, Modify, Model, Assess*) untuk memastikan tahapan analisis data berjalan secara sistematis dan terstruktur.

Untuk meningkatkan kualitas dataset, proses pelabelan dan validasi data dalam penelitian ini dilakukan secara manual dan sebagian dilakukan secara semi-otomatis. Untuk menguji konsistensi pelabelan, *Alpha Krippendorff* digunakan dan menemukan nilai 0,365, yang menunjukkan bahwa pelabelan memiliki keandalan yang memadai untuk pelatihan model. Teknik random oversampling digunakan untuk mengatasi ketidakseimbangan data. Hasil penelitian menunjukkan bahwa model *IndoBERT* yang telah disesuaikan (*fine-tuned*) memiliki nilai presisi, recall, dan skor F1 yang seimbang dengan akurasi sebesar 96%. Selain itu, analisis *confusion matrix* menunjukkan bahwa model tersebut memiliki kemampuan untuk menemukan pola sentimen dengan tingkat kesalahan yang sangat rendah.

Penggunaan model *IndoBERT* untuk analisis sentimen terhadap komentar Instagram yang berkaitan dengan kebijakan publik dan penerapan teknik *fine-tuning* untuk meningkatkan kinerja model adalah hal yang menyatukan kedua penelitian ini. Penelitian ini berfokus pada kebijakan Dewan Perdamaian, sedangkan penelitian sebelumnya berfokus pada kebijakan pajak pasar. Selain itu, penelitian ini tidak menggunakan pendekatan Kecerdasan Buatan yang Dapat Dijelaskan (XAI) seperti SHAP, untuk menjelaskan hasil prediksi model. Pendekatan ini dimaksudkan untuk meningkatkan keterbacaan dan transparansi analisis sentimen.

Penelitian keenam yang dilakukan oleh (Sumartha, 2025) dengan judul *"Implementasi IndoBERT untuk Analisis Sentimen Opini Publik terhadap Kebijakan Kenaikan UKT di Era Pemerintahan"* bertujuan untuk menganalisis sentimen publik terhadap kebijakan kenaikan Uang Kuliah Tunggal (UKT) menggunakan data dari media sosial Twitter. Penelitian ini mengimplementasikan model *IndoBERT* dengan tahapan *text preprocessing* yang komprehensif, meliputi *case folding*, *cleaning*, *tokenizing*, normalisasi, *stopword removal* dengan *exceptions list*, serta *stemming* untuk meningkatkan kualitas data sebelum proses pelatihan model.

Hasil penelitian menunjukkan bahwa model *IndoBERT* memahami konteks bahasa Indonesia dan mengklasifikasikan sentimen dengan akurasi 98%. Namun demikian,

penelitian ini menemukan beberapa masalah yang berkaitan dengan ketidakseimbangan kelas. Model ini mengidentifikasi sentimen positif dengan kesulitan (recall 0,29) tetapi berhasil dalam kelas negatif. Selain itu, dominasi sentimen negatif sebesar 87,9% dalam distribusi data menunjukkan bahwa opini publik mungkin menentang kenaikan biaya kuliah. Penelitian ini menemukan bahwa prapemrosesan yang disesuaikan dengan konteks linguistik memainkan peran yang signifikan dalam meningkatkan kinerja model.

Penggunaan model *IndoBERT* untuk analisis sentimen publik terhadap kebijakan pemerintah berdasarkan data media sosial adalah hal yang membuat penelitian ini dan penelitian sebelumnya memiliki kesamaan. Berbedanya terletak pada platform dan subjek penelitian, penelitian ini menggunakan data Instagram yang berkaitan dengan kebijakan Dewan Perdamaian, sedangkan penelitian (Sumartha, 2025) berfokus pada kenaikan biaya kuliah. Selain itu, penelitian (Sumartha, 2025) tidak menggunakan pendekatan Kecerdasan Buatan yang Dapat Dijelaskan (XAI), seperti SHAP, untuk menjelaskan hasil prediksi model. Pendekatan ini dimaksudkan untuk meningkatkan keterbacaan dan transparansi analisis sentimen.

Penelitian ketujuh yang dilakukan oleh (Asri et al., 2025) dengan judul *"Sentiment Analysis Based on Indonesian Language Lexicon and IndoBERT on User Reviews PLN Mobile Application"* bertujuan untuk mengevaluasi persepsi pengguna tentang aplikasi PLN Mobile berdasarkan ulasan yang diposting di platform Google Play Store. Dataset sebanyak 1.000 ulasan dikumpulkan melalui scraping web antara Januari dan Juni 2022 untuk studi ini. Untuk menghasilkan kumpulan data yang siap untuk dianalisis, pemrosesan data menggunakan pemrosesan teks sebelumnya.

Penelitian ini menggunakan pendekatan berbasis leksikon sentimen bahasa Indonesia, INSET (*Indonesia Sentiment Lexicon*) untuk melakukan pelabelan sentimen secara otomatis. Metode ini menghasilkan kumpulan data yang terdiri dari 482 ulasan positif, 144 ulasan negatif, dan 374 ulasan netral. Selanjutnya, untuk dataset pelatihan, pengujian, dan validasi, proses klasifikasi dilakukan menggunakan model *IndoBERT* dengan pembagian 80:10:10. Hasilnya menunjukkan bahwa model *IndoBERT* mencapai tingkat akurasi sebesar 81% dalam mengklasifikasikan sentimen pengguna terhadap aplikasi tersebut.

Penggunaan model *IndoBERT* untuk analisis sentimen dalam bahasa Indonesia dan

penerapan langkah-langkah prapemrosesan dan klasifikasi multi-kelas (positif, negatif, dan netral) adalah hal yang sama antara kedua penelitian ini. Sebagai objek analisis, kedua penelitian tersebut juga menggunakan data opini pengguna dari platform digital. Penelitian sebelumnya menggunakan ulasan aplikasi di Google Play Store dengan pelabelan berbasis leksikon (INSET) sedangkan penelitian ini menggunakan komentar Instagram yang berkaitan dengan kebijakan dewan perdamaian. Perbedaan terletak pada sumber data dan metode pelabelan. Selain itu, untuk meningkatkan keterbacaan hasil prediksi, penelitian ini menggunakan metode SHAP dan pendekatan *Explainable Artificial Intelligence* (XAI).

Tinjauan penelitian sebelumnya menunjukkan bahwa model berbasis *transformer*, khususnya *IndoBERT*, telah terbukti berhasil melakukan analisis sentimen pada teks berbahasa Indonesia, terutama pada data media sosial seperti Instagram, TikTok, YouTube, dan Twitter. Dibandingkan dengan metode pembelajaran mesin konvensional seperti *Support Vector Machines* (SVM) dan *Random Forest*, *IndoBERT* lebih baik dalam memahami konteks linguistik, termasuk variasi dalam bahasa informal. Selain itu, penelitian telah menunjukkan bahwa *IndoBERT* sangat akurat dan konsisten dalam mengkategorikan sentimen pada berbagai masalah kebijakan publik ke dalam kategori positif, negatif, atau netral.

Namun demikian, sebagian besar penelitian masih berfokus pada aspek performa model, seperti akurasi, *precision*, *recall*, dan *f1-score*, tanpa memberikan penjelasan yang memadai terkait proses pengambilan keputusan model. Hal ini menyebabkan model yang digunakan cenderung bersifat *black box* dan kurang transparan, sehingga menyulitkan pengguna dalam memahami kontribusi fitur atau kata tertentu terhadap hasil klasifikasi sentimen. Padahal, dalam konteks analisis opini publik, transparansi dan interpretabilitas menjadi aspek yang sangat penting untuk meningkatkan kepercayaan terhadap hasil analisis.

Oleh karena itu, tujuan dari penelitian ini adalah untuk mengatasi perbedaan ini dengan mengintegrasikan pendekatan Kecerdasan Buatan yang Dapat Dijelaskan *Explainable Artificial Intelligence* (XAI), khususnya metode SHAP (*SHapley Additive exPlanations*), ke dalam model *IndoBERT*. Tujuan dari integrasi ini adalah untuk meningkatkan keterbacaan hasil prediksi dengan meningkatkan kejelasan mereka. Selain itu, sebagai

objek analisis, penelitian ini menggunakan data komentar Instagram dari akun resmi pemerintah yang berkaitan dengan kebijakan Dewan Perdamaian. Tujuan dari penelitian ini adalah untuk membuat model yang efektif dan memberikan penjelasan yang jelas dan mudah dipahami tentang bagaimana kebijakan tersebut mempengaruhi persepsi publik.

2.2 Landasan Teori

2.2.1 Analisis Sentimen

Analisis sentimen adalah bagian dari pemrosesan bahasa alam (NLP) yang digunakan untuk menemukan, mengekstrak, dan mengkategorikan pendapat atau persepsi yang ada dalam teks ke dalam kategori tertentu, seperti positif, negatif, atau netral. Metode ini memungkinkan pemrosesan data teks untuk memahami tren opini publik tentang hal-hal atau masalah tertentu. Akibat volume data teks yang meningkat yang dihasilkan oleh media sosial, forum, dan platform digital lainnya, analisis sentimen telah menjadi sangat penting dalam dunia digital saat ini (Winarni, Achmad Hindasyah, 2025)

Media sosial telah berkembang menjadi sumber data utama untuk analisis sentimen seiring dengan pesatnya kemajuan teknologi informasi. Data dari media sosial beragam, tidak terorganisir, dan biasanya ditulis dalam bahasa informal. Akibatnya, analisis sentimen sangat penting untuk mengubah data teks ini menjadi informasi penting untuk pengambilan keputusan. Studi menunjukkan bahwa analisis sentimen dapat membantu memahami persepsi masyarakat terhadap fenomena atau kebijakan publik secara objektif dan sistematis (Setiawan Putra Mahardika, Rodhiyah Mardhiyyah, 2025).

Proses analisis sentimen biasanya terdiri dari beberapa tahap utama: pengumpulan data, prapemrosesan teks, pelabelan data, pemodelan, dan evaluasi kinerja model. Pada tahap prapemrosesan, kualitas data ditingkatkan dengan membersihkan teks, normalisasi kata, tokenisasi, dan penghapusan kata penghalang. Selanjutnya, data diklasifikasikan menggunakan algoritma khusus yang digunakan untuk mengidentifikasi polaritas sentimen dalam teks (Mursidah et al., 2025).

Analisis sentimen dapat dilakukan dengan berbagai cara, seperti pembelajaran mesin dan pembelajaran mendalam hingga pendekatan berbasis leksikon. Metode pembelajaran mesin seperti *Naive Bayes* dan *Support Vector Machines* (SVM) telah digunakan secara luas dalam klasifikasi teks. Namun, teknik ini terbatas pada

pemahaman konteks linguistik yang kompleks. Pendekatan *deep learning* terutama yang menggunakan model berbasis *transformer* seperti *IndoBERT* menawarkan solusi yang lebih baik karena mampu menangkap konteks semantik yang lebih luas (Anugerah et al., 2024).

Selain itu, analisis sentimen telah digunakan secara luas di berbagai bidang, seperti menilai pendapat konsumen tentang produk, menilai kebijakan pemerintah, dan melacak tanggapan masyarakat terhadap masalah sosial. Penelitian di berbagai bidang menunjukkan bahwa analisis sentimen dapat memberikan gambaran persepsi masyarakat dalam waktu nyata, yang dapat digunakan sebagai dasar pengambilan keputusan berbasis data (Putri et al., 2025).

Berdasarkan uraian tersebut, dapat disimpulkan bahwa analisis sentimen merupakan metode yang efektif untuk mengolah dan memahami opini publik dari data teks yang besar dan tidak terstruktur. Dengan memanfaatkan teknik NLP dan model pembelajaran mesin, analisis sentimen mampu memberikan insight yang bernilai dalam berbagai konteks, termasuk dalam memahami respons masyarakat terhadap kebijakan publik melalui media sosial.

2.2.2 *Board of Peace*

Board of Peace (BoP) atau Dewan Perdamaian merupakan badan internasional yang dibentuk untuk mengawal proses transisi, stabilisasi, rehabilitasi, dan rekonstruksi Gaza pascakonflik. Pembentukan badan ini bertujuan untuk mendukung terciptanya perdamaian yang berkelanjutan di Palestina melalui pengawasan terhadap administrasi, stabilisasi keamanan, serta pemulihan tata kelola sipil di wilayah Gaza (BPMI Setpres, 2026).

Menurut (BPMI Setpres, 2026), *Board of Peace* dibentuk sebagai bagian dari *Comprehensive Plan to End the Gaza Conflict (20-Point Roadmap)* dan memperoleh dukungan dari Dewan Keamanan Perserikatan Bangsa-Bangsa (PBB) melalui Resolusi 2803 Tahun 2025. Dalam pelaksanaannya, *Board of Peace* memiliki mandat untuk mengawasi pelaksanaan gencatan senjata, mendukung proses rekonstruksi Gaza, serta memastikan transisi menuju perdamaian yang berkelanjutan sesuai dengan hukum internasional dan prinsip solusi dua negara (*two-state solution*).

Sebagai negara yang menganut politik luar negeri bebas dan aktif, Indonesia memutuskan untuk bergabung dalam *Board of Peace* sebagai bentuk komitmen dalam mendukung perdamaian dan stabilitas internasional, khususnya bagi Palestina. Menteri Luar Negeri Republik Indonesia menyatakan bahwa keikutsertaan Indonesia didasarkan pada konsistensi pemerintah dalam mendukung penyelesaian konflik secara damai serta upaya rehabilitasi pascakonflik di Gaza. Selain itu, Indonesia memandang penting untuk terlibat secara langsung agar setiap kebijakan dan langkah yang diambil oleh *Board of Peace* tetap berorientasi pada kemerdekaan Palestina dan terwujudnya solusi dua negara (BPMI Setpres, 2026).

Keikutsertaan Indonesia dalam *Board of Peace* menjadi perhatian masyarakat dan memunculkan berbagai tanggapan di ruang publik. Sebagian masyarakat memandang kebijakan tersebut sebagai bentuk komitmen Indonesia dalam memperjuangkan perdamaian dunia, sedangkan sebagian lainnya menyampaikan kritik maupun keraguan terhadap dampak dan implikasinya bagi kepentingan nasional. Beragam opini tersebut banyak disampaikan melalui media sosial, termasuk Instagram, sehingga menghasilkan data yang dapat dimanfaatkan untuk mengidentifikasi kecenderungan sentimen masyarakat. Oleh karena itu, penelitian ini menggunakan komentar Instagram sebagai sumber data untuk menganalisis sentimen masyarakat terhadap keikutsertaan Indonesia dalam *Board of Peace* menggunakan model *IndoBERT* dan metode SHAP.

2.2.3 *Natural Language Processing* (NLP)

Aspek dari kecerdasan buatan (AI) yang berfokus pada interaksi antara komputer dan bahasa manusia adalah pemrosesan bahasa alami (NLP). NLP memungkinkan komputer untuk memahami, memproses, dan membuat bahasa alami. Ini membuatnya berguna untuk banyak hal, seperti analisis sentimen, penerjemahan bahasa, chatbot, sistem pencarian informasi, dan sebagainya. Teknologi memainkan peran penting dalam memfasilitasi komunikasi manusia-mesin yang lebih alami dan efektif (Salamun et al., 2024) .

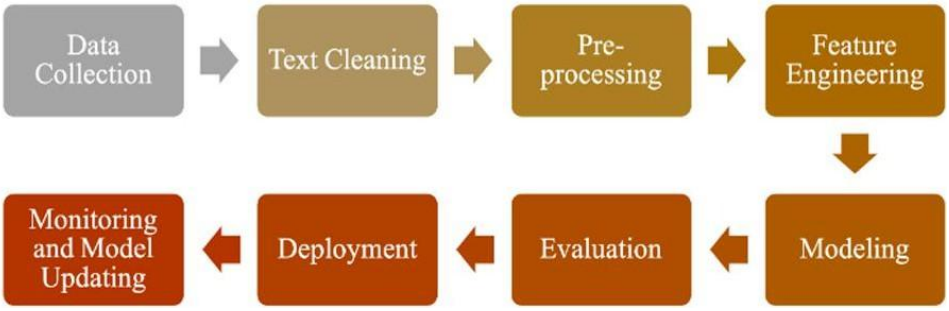
Secara umum, NLP bekerja dengan menggabungkan konsep linguistik dan komputasi untuk memproses teks atau bahasa alami menjadi bentuk yang dapat dipahami oleh mesin. Proses dalam NLP melibatkan beberapa tahapan utama, antara lain analisis leksikal, analisis sintaksis, analisis semantik, serta analisis pragmatik. Setiap tahapan tersebut

saling berkaitan dalam memahami makna bahasa secara menyeluruh. Analisis leksikal berfokus pada pembentukan kata, sedangkan analisis sintaksis mengkaji struktur kalimat. Selanjutnya, analisis semantik bertujuan memahami makna, dan analisis pragmatik mempertimbangkan konteks penggunaan Bahasa (Badawi, 2021)

Dalam implementasinya, NLP banyak digunakan untuk mengolah data teks yang tidak terstruktur, seperti komentar media sosial, berita, maupun dokumen digital lainnya. NLP memungkinkan proses ekstraksi informasi penting dari data teks secara otomatis serta membantu merangkum informasi sehingga lebih mudah dianalisis. Kemampuan tersebut menjadikan NLP banyak dimanfaatkan dalam berbagai penelitian berbasis teks untuk memperoleh informasi yang relevan dari kumpulan data yang besar (Siswipraptini et al., 2023). Penggunaan NLP memungkinkan ekstraksi informasi secara otomatis dari data dalam jumlah besar, sehingga dapat menghasilkan insight yang bernilai. Penelitian menunjukkan bahwa NLP efektif digunakan dalam analisis sentimen terhadap kebijakan publik, di mana sistem mampu mengelompokkan opini masyarakat berdasarkan polaritas sentimen secara akuratn (Iim Abdurrohim & Rahman, 2024)

Dengan kemajuan teknologi, proses pengolahan bahasa natural telah berkembang dari pendekatan berbasis aturan dan statistik menuju pendekatan pemahaman mendalam. Model berbasis pemahaman mendalam, terutama yang bergantung pada arsitektur transformasi seperti *BERT* dan *IndoBERT*, memiliki kapasitas untuk memahami konteks linguistik secara lebih mendalam dibandingkan metode sebelumnya. Ini membuat proses pengolahan bahasa natural semakin akurat dan fleksibel dalam menangani variasi linguistik, terutama dalam data media sosial yang dinamis dan informal.

Adapun tahapan dari NLP sebagai berikut.



Gambar 2. 1 Tahapan NLP

Sumber: (M. Rizky Suherlan, Asriyanik, 2023)

1. *Data Collection*

Tahap awal proses *natural linguistic* (NLP) berfokus pada pengumpulan data teks sebagai bahan utama untuk analisis. Data dapat diperoleh dari berbagai sumber tetapi tidak terbatas pada, basis data, dokumen digital, interaksi pengguna, atau platform lain yang sesuai dengan tujuan penelitian.

2. *Text Cleaning*

Pembersihan teks adalah prosedur yang digunakan untuk menghilangkan berbagai elemen teks yang tidak penting, yang menghasilkan data yang lebih terstruktur. Menghapus tanda baca, membuat huruf menjadi huruf kecil (*case folding*), menstandarkan kata-kata yang tidak baku, menghilangkan kata-kata *stopword*s, *stemming* untuk mengekstrak kata dasar, dan tokenisasi untuk memecah teks menjadi kata-kata atau frasa spesifik ini semua adalah langkah-langkah dalam proses ini.

3. *Pre-processing*

Prapemrosesan, langkah berikutnya setelah pembersihan teks, memiliki tujuan untuk mempersiapkan data secara lebih baik sebelum ekstraksi fitur. Pada tahap ini, teks dinormalisasi dan diproses untuk menghasilkan representasi yang lebih konsisten, yang siap untuk digunakan dalam analisis selanjutnya.

4. *Feature Engineering*

Pada tahap ini, teks yang telah diproses diubah ke dalam format numerik sehingga algoritma pembelajaran mesin dapat memahaminya. Tujuan dari proses ini adalah untuk mengekstrak fitur-fitur penting yang mewakili informasi dalam teks, sehingga proses pembelajaran model secara efektif dibantu.

5. *Modelling*

Fase pemodelan adalah proses pembuatan dan pelatihan model menggunakan data yang telah direpresentasikan dalam bentuk fitur. Selama fase ini, model mempelajari pola-pola dalam data untuk melakukan tugas-tugas tertentu, seperti klasifikasi atau prediksi, sehingga menghasilkan model yang siap digunakan.

6. *Evaluation*

Untuk mengukur kinerja model yang telah dibangun, evaluasi ini dilakukan. Tahap ini dilakukan untuk memastikan bahwa model dapat menghasilkan keluaran yang akurat dan memenuhi persyaratan untuk akurasi prediksi dan kualitas hasil.

7. *Deployment*

Proses mengintegrasikan model yang telah dikembangkan ke dalam lingkungan nyata sehingga dapat digunakan oleh pengguna disebut fase *deployment*. Ini akan diintegrasikan ke dalam sistem atau aplikasi tertentu sehingga dapat digunakan secara langsung.

8. *Monitoring and Model Updating*

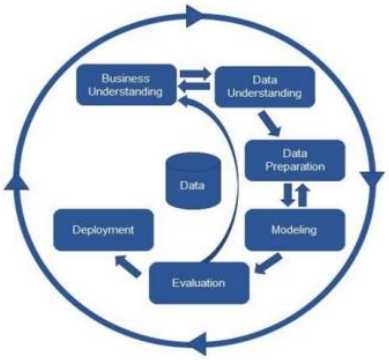
Setelah sistem diterapkan, fase ini dilakukan dengan memantau kinerja model secara berkala. Ini memastikan model tetap relevan dan mampu memberikan hasil yang optimal dengan memperbaruinya jika terjadi penurunan kinerja atau perubahan karakteristik data.

Penelitian ini berfokus pada pengembangan, evaluasi, dan penerapan proses linguistik natural (NLP). Tahapan NLP yang diterapkan termasuk pengumpulan data, prapemrosesan teks (termasuk *case folding*, pembersihan teks, dan normalisasi kata), pemodelan menggunakan *IndoBERT*, evaluasi model menggunakan *confusion matrix* dan laporan klasifikasi, dan interpretasi model menggunakan SHAP.

2.2.4 *Cross-Industry Standard Process for Data Mining* (CRISP-DM)

Salah satu metode yang paling umum digunakan dalam pengembangan proyek penambangan data dan ilmu data adalah CRISP-DM, proses standar industri untuk penambangan data. Metodologi ini menyediakan kerangka kerja sistematis yang membantu para peneliti mengelola proses analisis data secara terstruktur dari tahap pemahaman masalah hingga penerapan model. CRISP-DM fleksibel dan dapat digunakan dalam berbagai domain, seperti AI, pembelajaran mesin, dan analisis data (Rianti et al., 2023).

Kerangka kerja CRISPDM terdiri atas enam tahapan utama yang digambarkan sebagai berikut :



Gambar 2. 2Alur CRISP-DM

Sumber: (Rifai et al., 2019)

1. *Business Understanding*

Tahapan ini dilakukan untuk mendefinisikan permasalahan yang akan diteliti sekaligus merumuskan tujuan utama yang ingin dicapai. Pada fase ini, peneliti mengkaji kebutuhan, batasan masalah, serta sasaran penelitian agar proses analisis data yang dilakukan tetap terarah dan sesuai dengan konteks permasalahan.

2. *Data Understanding*

Pada tahap ini, fokusnya adalah proses eksplorasi diri dengan data yang digunakan dalam penelitian. Beberapa tindakan yang dilakukan termasuk pengumpulan data awal, eksplorasi isi data, analisis pola-pola yang muncul, dan pemeriksaan kualitas data untuk memastikan bahwa data tersebut layak untuk diproses lebih lanjut.

3. *Data Preparation*

Pada tahap ini, data diproses dan disiapkan untuk digunakan dalam proses pemodelan. Proses ini mencakup membersihkan data dari kesalahan atau gangguan, memilih data yang relevan, mengubah formatnya, normalisasi, dan pelabelan data sesuai kebutuhan untuk membuatnya lebih terstruktur dan siap digunakan.

4. *Modeling*

Pada tahap pemodelan, model dilatih menggunakan data yang telah disiapkan dan parameternya disesuaikan untuk mencapai kinerja model yang optimal.

5. *Evaluation*

Sejauh mana model yang dikembangkan dapat memenuhi tujuan penelitian adalah tujuan fase evaluasi serta untuk menilai apakah model sudah siap digunakan atau masih memerlukan penyempurnaan, metrik khusus digunakan untuk menilai jenis kasus yang diteliti.

6. *Deployment*

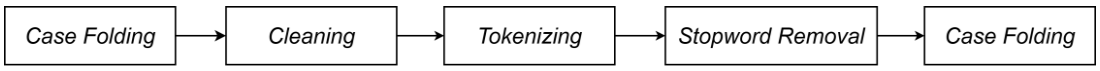
Laporan analisis, sistem yang siap digunakan, atau saran yang dapat digunakan sebagai dasar pengambilan keputusan adalah hasil dari fase penerapan, di mana hasil analisis atau model diterapkan ke lingkungan dunia nyata.

2.2.5 *Text Processing*

Pemrosesan teks merupakan tahap awal dalam pemrosesan data teks dalam bidang Pemrosesan Bahasa Alami (NLP), yang bertujuan untuk membersihkan dan mempersiapkan data agar dapat digunakan dalam analisis lebih lanjut. Tahap ini sangat penting karena kualitas data yang dihasilkan akan memengaruhi kinerja model, terutama dalam analisis sentimen.

Menurut (Hadiprakoso et al., 2023) text preprocessing merupakan tahap krusial yang melibatkan proses pembersihan teks seperti penghapusan simbol, emoji, dan *stopword*s, serta normalisasi teks seperti lowercasing, *stemming*, dan lemmatization. Proses ini terbukti dapat meningkatkan akurasi model, terutama pada analisis sentimen berbahasa Indonesia.

Tahapan dari *text processing* sebagai berikut.



Gambar 2. 3 Tahapan *Text Processing*

Sumber: (Hadiprakoso et al., 2023)

1. *Case folding*

Tahap *case folding* dilakukan untuk mengubah seluruh huruf dalam teks menjadi huruf kecil secara konsisten. Hal ini bertujuan agar perbedaan penulisan huruf kapital tidak menyebabkan sistem menganggap kata yang sama sebagai entitas yang berbeda, sehingga proses analisis menjadi lebih seragam.

2. *Cleaning*

Dalam tahap pembersihan, unsur-unsur teks yang tidak diperlukan, seperti angka, tanda baca, simbol, dan tautan URL, dihilangkan. Hal ini dilakukan karena unsur-unsur ini tidak memberikan kontribusi yang signifikan kepada makna teks selama proses analisis.

3. *Tokenizing*

Tokenisasi adalah proses yang memecah teks, yang semula terdiri dari kalimat atau paragraf, menjadi bagian yang lebih kecil, yaitu kata-kata. Ini memungkinkan teks yang tidak terstruktur dipecahkan menjadi bagian yang lebih sederhana untuk pemrosesan lebih lanjut.

4. *Stopword Removal*

Tujuan dari tahap *stopword* removal adalah untuk menyaring dan menghilangkan kata-kata yang sering muncul tetapi tidak penting (seperti "dan", "yang", "di", dan kata-kata serupa lainnya). Proses ini lebih efisien dan memfokuskan analisis pada kata-kata yang lebih relevan.

5. *Stemming*

Pada tahap *stemming*, kata-kata yang masih mengandung imbuhan dikembalikan ke bentuk dasarnya. Proses ini umumnya dilakukan menggunakan Sastrawi, yang berfungsi menghapus awalan, sisipan, maupun akhiran sehingga variasi bentuk suatu kata dapat diseragamkan menjadi satu bentuk dasar.

Secara umum, text preprocessing dapat mencakup berbagai tahapan, *seperti case folding, cleaning, tokenizing, stopword removal*, dan *stemming*. Namun, tahapan yang diterapkan pada setiap penelitian dapat berbeda bergantung pada karakteristik data dan model yang digunakan. Pada penelitian ini, preprocessing yang dilakukan meliputi *case folding, cleaning*, dan *normalization.* Proses tokenisasi tidak dilakukan secara manual karena telah ditangani oleh *BERT Tokenizer* pada model *IndoBERT,* sedangkan *stopword removal* dan *stemming* tidak diterapkan agar informasi kontekstual dalam teks tetap terjaga selama proses *fine-tuning.*

2.2.6 *Deep Learning*

Sebagai salah satu pendekatan dalam *machine learning, Deep learning* menggunakan

arsitektur jaringan saraf tiruan yang terdiri atas banyak lapisan untuk mempelajari pola dan representasi data secara otomatis. Pendekatan ini mampu mengekstraksi informasi penting dari data dalam jumlah besar tanpa bergantung pada proses ekstraksi fitur secara manual, sehingga banyak dimanfaatkan dalam aplikasi pengolahan citra, pengenalan suara, maupun *Natural Language Processing* (NLP).

Menurut penelitian oleh (No et al., 2025) *deep learning* merupakan pengembangan dari neural network yang memiliki banyak lapisan tersembunyi (*hidden layer*) sehingga mampu menghasilkan performa prediksi yang lebih tinggi dibandingkan metode konvensional.

2.2.7 Arsitektur *Transformer* dan Model *BERT*

Arsitektur *Transformer* merupakan model *deep learning* yang dikembangkan untuk mengolah data berbentuk sekuens tanpa menggunakan mekanisme rekuren (*recurrent*) maupun konvolusi (*convolution*). Sebagai gantinya, model ini memanfaatkan mekanisme attention untuk menangkap hubungan antar kata dalam sebuah kalimat, sehingga mampu memahami konteks secara lebih efektif dibandingkan pendekatan sebelumnya (Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, 2023).

Secara umum, *Transformer* terdiri atas dua komponen utama, yaitu *encoder* dan *decoder*. *Encoder* bertugas mengubah teks masukan menjadi representasi vektor yang mengandung informasi kontekstual, sedangkan *decoder* memanfaatkan representasi tersebut untuk menghasilkan keluaran yang sesuai. Kedua komponen dibangun dari beberapa lapisan identik yang memungkinkan model mempelajari pola bahasa secara mendalam (Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, 2023).

Setiap lapisan pada *encoder* tersusun atas *multi-head self-attention* dan *feed forward neural network*. Melalui mekanisme *self-attention*, model dapat mengukur tingkat keterkaitan antar kata dalam satu kalimat sehingga setiap kata diproses dengan mempertimbangkan konteks di sekitarnya. Proses ini melibatkan tiga elemen utama, yaitu *query, key*, dan *value*, yang digunakan untuk menghitung bobot attention dan membentuk representasi kontekstual dari setiap kata.

Selain *self-attention, Transformer* juga menerapkan *multi-head attention*, yaitu

mekanisme yang memungkinkan beberapa proses attention dijalankan secara bersamaan. Pendekatan ini membuat model mampu menangkap berbagai jenis hubungan antar kata dari beragam perspektif, sehingga representasi teks yang dihasilkan menjadi lebih kaya dan informatif (Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, 2023).

Karena *Transformer* tidak menggunakan urutan sekuensial seperti pada model RNN, maka informasi posisi kata ditambahkan melalui teknik *positional encoding*. Teknik ini bertujuan untuk memberikan informasi terkait urutan kata dalam kalimat sehingga model tetap dapat memahami struktur bahasa dengan baik(Ashish Vaswani et al., 2023).

BERT (*Bidirectional Encoder Representations from Transformer*s) merupakan model berbasis *Transformer* yang dikembangkan untuk memahami konteks bahasa secara dua arah (*bidirectional*). Berbeda dengan model sebelumnya yang memproses teks secara satu arah, *BERT* memanfaatkan informasi dari kata-kata sebelum dan sesudah suatu token sehingga mampu menghasilkan representasi konteks yang lebih komprehensif (Devlin et al., 2019).

Dalam arsitekturnya, BERT hanya menggunakan bagian *encoder* dari *Transformer*. Selama proses pre-training, model dilatih menggunakan dua tugas utama, yaitu *Masked Language Model* (MLM) dan *Next Sentence Prediction* (NSP). Pada MLM, sejumlah token dalam kalimat disembunyikan agar model dapat memprediksi token yang hilang berdasarkan konteks di sekitarnya. Sementara itu, NSP digunakan untuk mempelajari hubungan antar kalimat dengan menentukan apakah dua kalimat saling berurutan atau tidak (Devlin et al., 2019).

Setelah tahap pre-training selesai, BERT dapat diadaptasi ke berbagai tugas *Natural Language Processing* (NLP) melalui proses *fine-tuning*. Pada tahap ini, model dilatih kembali menggunakan dataset yang sesuai dengan tugas tertentu, seperti klasifikasi teks, analisis sentimen, dan pengenalan entitas bernama, dengan menambahkan layer output sesuai kebutuhan (Devlin et al., 2019).

Keunggulan utama dari *Transformer* dan BERT terletak pada kemampuannya dalam menangkap hubungan konteks jangka panjang serta memproses data secara paralel. Hal ini membuat kedua model tersebut menjadi sangat efektif dan banyak digunakan dalam berbagai penelitian serta aplikasi NLP modern.

2.2.8 *Hyperparameter* Optimization

Dalam pengembangan model *machine learning* dan *deep learning*, pemilihan *Hyperparameter* merupakan salah satu faktor penting yang memengaruhi performa model. *Hyperparameter* adalah parameter yang nilainya ditentukan sebelum proses pelatihan dimulai dan tidak dipelajari secara otomatis selama proses pembelajaran model. Berbeda dengan parameter model yang diperbarui melalui proses pelatihan, *Hyperparameter* digunakan untuk mengatur proses pembelajaran sehingga model dapat menghasilkan performa yang optimal. Menurut (Bischl et al., 2023), pemilihan *Hyperparameter* yang kurang tepat dapat menyebabkan penurunan performa model serta meningkatkan risiko terjadinya *underfitting* maupun *overfitting*.

Proses pencarian kombinasi *Hyperparameter* yang optimal dikenal sebagai *Hyperparameter optimization* atau *Hyperparameter tuning*. Tujuan dari proses ini adalah memperoleh konfigurasi parameter yang mampu menghasilkan performa terbaik pada model berdasarkan metrik evaluasi tertentu. *Hyperparameter* optimization menjadi penting karena setiap dataset memiliki karakteristik yang berbeda sehingga tidak terdapat satu kombinasi *Hyperparameter* yang selalu memberikan hasil terbaik pada seluruh kasus. Oleh karena itu, proses optimasi diperlukan untuk menyesuaikan konfigurasi model dengan karakteristik data yang digunakan.

Pada model berbasis *Transformer* seperti BERT dan *IndoBERT*, beberapa *Hyperparameter* yang umum digunakan meliputi *batch size*, *learning rate*, *epoch*, dan *weight decay*. Keempat *Hyperparameter* tersebut memiliki pengaruh yang signifikan terhadap proses pelatihan model dan kualitas hasil prediksi yang dihasilkan.

Batch size, learning rate, epoch, dan weight decay merupakan *Hyperparameter* yang berperan penting dalam proses pelatihan model. Pemilihan nilai yang tepat akan memengaruhi kecepatan pembelajaran, stabilitas proses pelatihan, serta kemampuan model dalam melakukan generalisasi terhadap data baru.

Batch size menentukan banyaknya data yang diproses dalam satu iterasi pelatihan. Penggunaan batch size yang terlalu kecil dapat menyebabkan pembaruan parameter menjadi kurang stabil, sedangkan nilai yang terlalu besar membutuhkan sumber daya komputasi yang lebih tinggi. Pada proses *fine-tuning* BERT, (Devlin et al., 2019) merekomendasikan penggunaan batch size pada rentang 16–32.

Learning rate mengatur besar kecilnya langkah pembaruan bobot model selama proses optimasi. Nilai yang terlalu tinggi berisiko membuat proses pelatihan tidak stabil, sedangkan nilai yang terlalu rendah menyebabkan model memerlukan waktu lebih lama untuk mencapai konvergensi. Menurut (Devlin et al., 2019) *learning rate* yang umum digunakan pada *fine-tuning* BERT berada pada kisaran 2e-5 hingga 5e-5.

Epoch menunjukkan jumlah pengulangan proses pelatihan terhadap seluruh data training. Penambahan jumlah epoch memberikan kesempatan bagi model untuk mempelajari pola data secara lebih mendalam. Namun, penggunaan epoch yang berlebihan dapat meningkatkan risiko *overfitting*, yaitu kondisi ketika model terlalu menyesuaikan diri dengan data pelatihan sehingga kemampuan generalisasinya terhadap data baru menurun.

Selain ketiga *Hyperparameter* tersebut, *weight decay* berfungsi sebagai teknik regularisasi untuk membatasi kompleksitas model. Parameter ini memberikan penalti terhadap bobot yang terlalu besar sehingga membantu mengurangi *overfitting* dan meningkatkan kemampuan model dalam melakukan generalisasi pada data yang belum pernah dipelajari.

Karakteristik setiap dataset yang berbeda menyebabkan tidak ada satu kombinasi *Hyperparameter* yang selalu memberikan hasil terbaik. Oleh karena itu, diperlukan proses optimasi *Hyperparameter* untuk memperoleh konfigurasi yang paling sesuai dengan data yang digunakan.

2.2.8.1 Optuna

Optuna merupakan framework *open-source* yang dikembangkan untuk mengoptimalkan *Hyperparameter* pada model *machine learning* maupun *deep learning*. Framework ini diperkenalkan oleh (Akiba et al., 2019) pada tahun 2019 melalui publikasi berjudul *Optuna: A Next-generation Hyperparameter Optimization Framework* dalam *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* Pengembangan Optuna bertujuan untuk mengotomatisasi proses pencarian *Hyperparameter*, sehingga proses optimasi menjadi lebih efisien dibandingkan penentuan parameter secara manual.

Dalam proses optimasi, Optuna menjalankan sejumlah trial dengan kombinasi *Hyperparameter* yang berbeda pada setiap percobaan. Performa model kemudian

dievaluasi untuk setiap trial, sehingga konfigurasi yang memberikan hasil terbaik dapat dipilih sebagai *Hyperparameter* optimal. Pendekatan ini memungkinkan proses pencarian parameter dilakukan secara sistematis serta menghasilkan konfigurasi yang lebih sesuai dengan karakteristik dataset yang digunakan.

Salah satu keunggulan utama Optuna adalah penggunaan konsep *define-by-run*, yaitu metode yang memungkinkan ruang pencarian (*search space*) *Hyperparameter* dibangun secara dinamis selama program dijalankan. Pendekatan ini memberikan fleksibilitas yang lebih tinggi dibandingkan metode optimasi konvensional karena pengguna dapat menentukan rentang pencarian *Hyperparameter* secara lebih mudah dan menyesuaikannya dengan kebutuhan penelitian (Akiba et al., 2019) .

Dalam Optuna, proses optimasi dilakukan melalui dua komponen utama, yaitu *study* dan *trial*. *Study* merupakan keseluruhan proses optimasi yang dilakukan untuk mencari kombinasi *Hyperparameter* terbaik, sedangkan *trial* merupakan satu kali percobaan yang berisi kombinasi *Hyperparameter* tertentu. Pada setiap *trial*, model akan dilatih dan dievaluasi menggunakan fungsi objektif (*objective function*) yang telah ditentukan. Hasil evaluasi tersebut kemudian digunakan oleh Optuna untuk menentukan kombinasi *Hyperparameter* yang menghasilkan performa terbaik.

Pada penelitian ini, Optuna digunakan untuk melakukan optimasi *Hyperparameter* pada proses *fine-tuning* model *IndoBERT*. *Hyperparameter* yang dioptimasi meliputi *batch size*, *learning rate*, *epoch*, dan *weight decay*. Proses optimasi dilakukan sebelum pelatihan model akhir dengan tujuan memperoleh kombinasi *Hyperparameter* yang mampu menghasilkan performa klasifikasi sentimen yang optimal. *Hyperparameter* terbaik yang diperoleh dari proses optimasi selanjutnya digunakan pada tahap *fine-tuning* model *IndoBERT* untuk melakukan klasifikasi sentimen masyarakat terhadap keikutsertaan Indonesia dalam kebijakan *Board of Peace* berdasarkan komentar Instagram.

2.2.9 *IndoBERT*

IndoBERT adalah model bahasa yang telah dilatih sebelumnya berdasarkan arsitektur *Transformer* yang dikhususkan untuk memahami bahasa Indonesia. Ini adalah adaptasi dari BERT (Bidirectional Encoder Representations from *Transformer*s), yang dilatih menggunakan korpus bahasa Indonesia yang luas, sehingga lebih baik memahami konteks

linguistik daripada model umum atau multibahasa. Penelitian *"IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP"* oleh (Koto et al., 2020) memperkenalkan model *IndoBERT*. Makalah ini menawarkan kumpulan data tolok ukur dan model bahasa untuk berbagai tugas pengolahan bahasa natural.

IndoBERT bekerja dengan pendekatan bidirectional context understanding, yaitu membaca teks dari dua arah (kiri ke kanan dan kanan ke kiri) secara simultan. Pendekatan ini memungkinkan model memahami makna kata berdasarkan konteks kalimat secara keseluruhan, sehingga lebih efektif dalam menangani ambiguitas bahasa, terutama pada teks informal seperti media sosial. Kemampuan ini menjadikan *IndoBERT* sangat sesuai untuk tugas seperti analisis sentimen, klasifikasi teks, dan ekstraksi informasi dalam bahasa Indonesia (Setiawan Putra Mahardika, Rodhiyah Mardhiyyah, 2025).

Dalam implementasinya, *IndoBERT* biasanya digunakan melalui proses *fine-tuning*, yaitu menyesuaikan model yang telah dilatih sebelumnya dengan dataset spesifik sesuai kebutuhan penelitian. Proses ini memungkinkan model untuk beradaptasi dengan domain tertentu, seperti opini publik, ulasan pengguna, atau komentar media sosial. Berbagai penelitian menunjukkan bahwa *IndoBERT* mampu memberikan performa yang tinggi dalam klasifikasi sentimen, bahkan lebih unggul dibandingkan metode *deep learning* konvensional seperti LSTM (Purwanto, 2026).

Selain itu, *IndoBERT* juga terbukti efektif dalam menangani karakteristik bahasa Indonesia yang kompleks, seperti penggunaan bahasa tidak baku, singkatan, serta variasi gaya bahasa pada media sosial. Hal ini didukung oleh penelitian yang menunjukkan bahwa *IndoBERT* mampu mengklasifikasikan sentimen publik secara akurat pada berbagai isu, mulai dari kebijakan pemerintah hingga ulasan produk digital (Merdiansah et al., 2024).

Lebih lanjut, *IndoBERT* telah banyak digunakan dalam berbagai studi berbasis data sosial media di Indonesia, seperti analisis sentimen terhadap kebijakan publik, e-commerce, maupun layanan digital. Model ini memanfaatkan representasi embedding berbasis subword sehingga mampu mengenali kata-kata yang tidak umum atau mengalami perubahan bentuk, yang sering muncul dalam komunikasi digital masyarakat Indonesia (Putri et al., 2025).

Tahapan kerja *IndoBERT* dalam penelitian ini dimulai dari input berupa teks komentar

yang telah melalui proses *preprocessing*. Teks tersebut kemudian diproses menggunakan *BERT Tokenizer* untuk diubah menjadi token dan representasi numerik yang dapat dipahami oleh model. Setelah itu, token diproses oleh model *IndoBERT* untuk menghasilkan representasi kontekstual. Representasi tersebut kemudian digunakan oleh lapisan klasifikasi untuk menentukan kategori sentimen dari komentar, yaitu positif, negatif, atau netral.

Dengan berbagai keunggulan tersebut, *IndoBERT* menjadi salah satu model yang sangat relevan dalam penelitian ini, khususnya dalam menganalisis sentimen komentar pengguna media sosial. Kemampuannya dalam memahami konteks bahasa secara mendalam serta performa yang tinggi dalam berbagai penelitian menjadikan *IndoBERT* sebagai pendekatan yang tepat untuk memperoleh hasil analisis yang akurat dan representatif terhadap opini publik.

2.2.8.1 BERT Tokenizer pada *IndoBERT*

BERT Tokenizer merupakan komponen penting dalam pemrosesan teks sebelum data dimasukkan ke dalam model *IndoBERT*. Tokenizer berfungsi untuk mengubah teks mentah menjadi format numerik yang dapat diproses oleh model. Dalam implementasi model berbasis BERT, teks tidak langsung dimasukkan ke dalam model dalam bentuk kalimat utuh, melainkan terlebih dahulu dipecah menjadi token, dikonversi menjadi ID token, kemudian dilengkapi dengan informasi tambahan seperti attention mask. Dokumentasi Hugging Face menjelaskan bahwa tokenizer bertugas menyiapkan input untuk model, termasuk memecah teks menjadi token atau subword, mengubah token menjadi ID numerik, serta mengelola special token yang dibutuhkan oleh model (Hugging Face, 2025).

Tokenisasi pada BERT menggunakan pendekatan subword tokenization, yaitu memecah kata menjadi unit yang lebih kecil apabila kata tersebut tidak ditemukan secara utuh dalam kosakata model. Salah satu metode subword tokenization yang digunakan dalam BERT adalah WordPiece Tokenization. (Song et al., 2021) menjelaskan bahwa WordPiece merupakan metode tokenisasi yang digunakan pada BERT dan bekerja dengan strategi longest-match-first atau maximum matching, yaitu memilih potongan subword terpanjang yang terdapat dalam vocabulary.

Penggunaan subword tokenization bermanfaat dalam menangani kata yang jarang

muncul atau tidak tersedia dalam vocabulary model. Pada komentar media sosial, pengguna sering menulis kata tidak baku, singkatan, atau variasi penulisan tertentu. Dengan pendekatan subword, kata yang tidak dikenali secara utuh masih dapat direpresentasikan melalui potongan subword yang lebih kecil. Hal ini membantu model tetap dapat memproses teks informal, meskipun bentuk kata yang digunakan tidak selalu baku.

Dalam proses tokenisasi BERT, tokenizer menambahkan token khusus atau special token sesuai kebutuhan model. Pada tugas klasifikasi teks, token [CLS] umumnya diletakkan di awal input sebagai penanda representasi keseluruhan kalimat, sedangkan token [SEP] digunakan sebagai penanda akhir kalimat atau pemisah antar kalimat. Selain itu, terdapat token [PAD] untuk menyamakan panjang input, [UNK] untuk token yang tidak dikenali, dan [MASK] yang digunakan pada proses pra-latih BERT. Pengelolaan special token merupakan salah satu fungsi utama tokenizer dalam model berbasis *Transformer*(Hugging Face, 2025).

Hasil dari proses BERT Tokenizer biasanya berupa beberapa komponen utama, yaitu input_ids dan attention_mask. input_ids merupakan representasi numerik dari token yang telah disesuaikan dengan vocabulary model, sedangkan attention_mask digunakan untuk membedakan token asli dengan token padding. Token padding diperlukan karena model *Transformer* memproses data dalam bentuk batch dengan panjang input yang seragam. Hugging Face menjelaskan bahwa tokenizer mengubah teks menjadi array angka sebagai input model dan mengembalikan attention mask untuk menunjukkan token mana yang perlu diperhatikan oleh model(Hugging Face, 2025).

Dalam penelitian ini, BERT Tokenizer digunakan setelah tahap preprocessing teks, yaitu setelah komentar melalui proses *case folding* dan normalization. Penelitian ini tidak menggunakan tokenizing manual karena proses tokenisasi telah dilakukan oleh tokenizer bawaan *IndoBERT*. Tokenizing manual seperti memisahkan kata berdasarkan spasi tidak digunakan karena dapat menyebabkan hasil token tidak sesuai dengan vocabulary yang digunakan oleh model pra-latih. Tokenizer yang digunakan harus sesuai dengan model *IndoBERT* agar representasi token yang dihasilkan dapat diproses dengan benar oleh model. Dokumentasi Hugging Face juga menekankan bahwa vocabulary tokenizer harus sesuai dengan vocabulary model pra-latih yang digunakan(Hugging Face, 2025).

Secara umum, alur kerja BERT Tokenizer pada *IndoBERT* dalam penelitian ini adalah sebagai berikut:

1. Teks komentar hasil preprocessing dimasukkan ke tokenizer.
2. Tokenizer memecah teks menjadi token atau subword.
3. Tokenizer menambahkan special token seperti [CLS] dan [SEP].
4. Token dikonversi menjadi input_ids berdasarkan vocabulary *IndoBERT*.
5. Tokenizer membentuk attention_mask untuk membedakan token asli dan token padding.
6. Output tokenizer dimasukkan ke model *IndoBERT* untuk proses klasifikasi sentimen.

Dengan demikian, BERT Tokenizer berperan sebagai penghubung antara data teks dan model *IndoBERT*. Tokenizer memastikan bahwa teks komentar Instagram yang awalnya berbentuk bahasa alami dapat diubah menjadi representasi numerik yang sesuai dengan kebutuhan model. Oleh karena itu, meskipun penelitian ini tidak menggunakan tokenizing manual pada tahap preprocessing, proses tokenisasi tetap dilakukan secara otomatis melalui BERT Tokenizer bawaan *IndoBERT*.

2.2.10 XAI Berbasis SHAP

Explainable Artificial Intelligence (XAI) merupakan pendekatan dalam kecerdasan buatan yang bertujuan untuk meningkatkan transparansi dan interpretabilitas model *machine learning*, khususnya pada model kompleks seperti *deep learning* yang cenderung bersifat black box. Dalam konteks analisis sentimen, interpretabilitas menjadi aspek penting karena hasil prediksi tidak hanya perlu akurat, tetapi juga harus dapat dijelaskan secara logis agar dapat dipercaya oleh pengguna maupun pengambil keputusan (Mas'ud et al., 2025).

Seiring berkembangnya teknologi, kebutuhan terhadap model yang dapat dijelaskan semakin meningkat, terutama pada aplikasi yang berkaitan dengan opini publik, kebijakan, maupun sistem pendukung keputusan. XAI hadir sebagai solusi untuk memberikan pemahaman mengenai bagaimana suatu model menghasilkan prediksi tertentu, baik melalui pendekatan lokal (penjelasan pada satu data) maupun global (penjelasan keseluruhan model). Dengan demikian, XAI tidak hanya meningkatkan transparansi, tetapi juga membantu dalam proses validasi model serta mengurangi potensi

bias (Mas'ud et al., 2025).

Salah satu metode yang populer dalam XAI adalah SHAP (*SHapley Additive exPlanations*). SHAP merupakan metode interpretasi berbasis teori permainan (game theory), khususnya konsep Shapley Value, yang digunakan untuk menghitung kontribusi setiap fitur terhadap hasil prediksi model. Metode ini mampu memberikan nilai kontribusi yang adil untuk setiap fitur dengan mempertimbangkan semua kemungkinan kombinasi fitur dalam model (Pasaribu et al., 2026).

Secara matematis, kontribusi setiap fitur dalam SHAP dihitung menggunakan konsep Shapley Value yang dirumuskan sebagai berikut (Lundberg & Lee, 2017):

$$\phi_i = \sum_{\{S \subseteq N \setminus \{i\}\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} \left[f(S \cup \{i\}) - f(S)\right]$$

Keterangan:

- ϕ_i= nilai Shapley dari fitur ke-i
- S= himpunan subset fitur yang tidak memuat fitur ke-i
- N= seluruh fitur dalam model
- $f(S)$= output model menggunakan subset fitur S
- $f(S \cup \{i\})$= output model setelah fitur ke-i ditambahkan
- $|\ S\ |$= jumlah fitur pada subset S
- $|\ N\ |$= jumlah seluruh fitur

Keunggulan utama SHAP terletak pada kemampuannya dalam memberikan interpretasi baik secara lokal maupun global. Secara lokal, SHAP dapat menjelaskan alasan di balik satu prediksi tertentu, misalnya mengapa suatu komentar diklasifikasikan sebagai sentimen negatif. Sedangkan secara global, SHAP dapat menunjukkan fitur mana yang paling berpengaruh terhadap keseluruhan model. Hal ini menjadikan SHAP sebagai metode yang sangat efektif dalam memahami perilaku model *machine learning* secara menyeluruh (Pasaribu et al., 2026).

Dalam implementasinya, SHAP menyediakan berbagai bentuk visualisasi yang dapat disesuaikan dengan tujuan interpretasi, seperti bar plot, summary plot, force plot, waterfall plot, dan text plot. Setiap visualisasi memiliki fungsi yang berbeda dalam

menjelaskan perilaku model maupun kontribusi fitur terhadap hasil prediksi. Pada penelitian ini, visualisasi SHAP yang digunakan adalah global bar plot untuk menganalisis kontribusi fitur secara keseluruhan serta text plot untuk menjelaskan kontribusi setiap token pada prediksi suatu komentar.

Beberapa penelitian di Indonesia telah menunjukkan efektivitas penggunaan SHAP dalam meningkatkan interpretabilitas model. Penelitian oleh (Bukhori et al., 2025) berjudul *"GWO-Enhanced Hybrid Deep learning with SHAP for Explainable TLKM.JK Stock Forecasting"* menunjukkan bahwa integrasi SHAP mampu mengidentifikasi fitur penting seperti volume dan indikator teknis lainnya dalam prediksi saham, sekaligus memberikan interpretasi yang jelas terhadap hasil model.

Selain itu, penelitian oleh (Pasaribu et al., 2026) dengan judul *"Explainable Machine learning Models SHAP-based for Feature Importance Affecting Stunting Prevalence"* juga membuktikan bahwa SHAP dapat digunakan untuk mengidentifikasi faktor-faktor penting dalam model prediksi, sehingga hasil analisis menjadi lebih transparan dan mudah dipahami oleh pengguna non-teknis.

Penelitian lain oleh (Sakmar et al., 2026) berjudul *"An Explainable Spatio-Temporal Decision Support System Using XGBoost and SHAP for Urban Complaint Trend Prediction"* menunjukkan bahwa penggunaan SHAP dalam sistem berbasis XGBoost mampu meningkatkan kejelasan dalam memahami pola keluhan masyarakat secara spasial dan temporal, sehingga dapat mendukung pengambilan keputusan yang akurat

Berdasarkan diskusi di atas, dapat disimpulkan bahwa SHAP adalah metode AI yang efektif untuk meningkatkan transparansi dan keterbacaan model pembelajaran mesin. Dalam penelitian ini, SHAP digunakan untuk menjelaskan kontribusi setiap kata terhadap hasil prediksi model *IndoBERT* melalui interpretasi lokal dan global, memastikan proses klasifikasi sentimen berjalan dengan baik dan menjelaskan mengapa model pembelajaran mesin

2.2.11 Instagram

Instagram adalah platform media sosial berbasis visual yang memungkinkan pengguna berbagi foto dan video dan berinteraksi melalui fitur seperti suka, komentar, dan pesan langsung. Selain itu, platform ini berfungsi sebagai media untuk membangun jaringan sosial dan menyebarkan informasi secara luas di internet (Arlena, 2021). Instagram telah

berkembang menjadi salah satu platform media sosial paling populer berkat kemampuan untuk memfasilitasi interaksi real-time dan menampilkan konten yang menarik (Sazali & Sukriah, 2021).

Selain sebagai media komunikasi, Instagram berperan dalam membentuk opini publik melalui interaksi pengguna berupa komentar, like, dan share. Aktivitas tersebut mencerminkan respon dan persepsi masyarakat terhadap suatu isu, sehingga menjadikan Instagram sebagai sumber data yang relevan dalam penelitian analisis sentimen. Dengan karakteristik konten yang interaktif dan dinamis, Instagram mampu merepresentasikan opini publik secara langsung di lingkungan digital (Elfariani et al., 2024).

2.2.12 Jupyter Notebook

Jupyter Notebook merupakan aplikasi berbasis web yang digunakan untuk menulis, menjalankan, dan mendokumentasikan kode program secara interaktif. Jupyter Notebook banyak digunakan dalam bidang data science, *machine learning*, artificial intelligence, serta penelitian yang melibatkan pengolahan dan analisis data karena mampu mengintegrasikan kode program, teks dokumentasi, visualisasi data, dan hasil eksekusi program dalam satu lingkungan kerja yang sama.

Pada Jupyter Notebook, kode program dijalankan dalam bentuk *cell* sehingga proses pengembangan dan pengujian program dapat dilakukan secara bertahap. Fitur ini memudahkan peneliti dalam melakukan eksplorasi data, preprocessing, visualisasi, pelatihan model, serta evaluasi hasil tanpa harus menjalankan seluruh program secara berulang. Selain itu, Jupyter Notebook mendukung berbagai library Python yang umum digunakan dalam penelitian *machine learning*, seperti Pandas, NumPy, Matplotlib, Scikit-learn, PyTorch, dan *Transformer*s.

Dalam penelitian ini, Jupyter Notebook digunakan sebagai lingkungan pengembangan untuk seluruh proses analisis sentimen, mulai dari proses pemuatan dataset, pembersihan data, preprocessing teks, tokenisasi menggunakan BERT Tokenizer, optimasi *Hyperparameter* menggunakan Optuna, *fine-tuning* model *IndoBERT*, evaluasi model, hingga interpretasi hasil menggunakan SHAP. Penggunaan Jupyter Notebook memudahkan proses eksperimen karena setiap tahapan penelitian dapat dijalankan, diamati, dan didokumentasikan secara sistematis dalam satu file notebook.

Menurut (Rahman et al., 2023), Jupyter Notebook merupakan editor Python berbasis web yang dapat dijalankan melalui peramban (*browser*) dan digunakan untuk mendukung proses pemrograman secara interaktif. Selain itu, (Muhammad Radhi, 2022) menjelaskan bahwa Jupyter Notebook banyak dimanfaatkan dalam proses *exploratory data analysis* (EDA), visualisasi data, dan pengolahan data karena mampu mengintegrasikan proses analisis dan visualisasi dalam satu lingkungan kerja. Penggunaan Jupyter Notebook juga banyak diterapkan dalam penelitian *machine learning* karena mendukung proses eksperimen dan pengembangan model secara lebih fleksibel.

Dengan berbagai kemampuan tersebut, Jupyter Notebook menjadi salah satu perangkat lunak yang banyak digunakan dalam penelitian berbasis *machine learning* dan *Natural Language Processing* (NLP), termasuk pada penelitian analisis sentimen menggunakan model *IndoBERT*.

2.2.13 Metrik Evaluasi(*Confusion matrix*)

Evaluasi model adalah proses mengukur kemampuan model pembelajaran mesin untuk membuat prediksi yang sesuai dengan data aktual. Dalam tugas klasifikasi seperti analisis sentimen, evaluasi dilakukan dengan membandingkan hasil prediksi dengan label aktual untuk menentukan keandalan dan akurasi model.

Salah satu metode evaluasi yang umum digunakan adalah *confusion matrix*, yang terdiri dari empat komponen utama, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Komponen-komponen ini menjadi dasar dalam perhitungan berbagai metrik evaluasi(Pratiwi et al., 2020) .

Salah satu metrik evaluasi yang umum digunakan adalah akurasi, presisi, recall, dan skor F1. Akurasi menunjukkan proporsi prediksi yang benar, sedangkan presisi mengukur keakuratan prediksi positif. Recall digunakan untuk menilai kemampuan model untuk menemukan semua titik data positif, dan skor F1 adalah kombinasi dari presisi dan recall yang memberikan evaluasi seimbang terhadap model (Pratiwi et al., 2020).

1. Akurasi(*Accuracy*)

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

Akurasi merupakan perbandingan antara total prediksi yang tepat dengan keseluruhan jumlah data yang diuji. Metrik ini digunakan untuk menilai kemampuan sistem dalam

menghasilkan prediksi yang benar secara umum.

2. Presisi(*Precision*)

$$Precision = \frac{(TP)}{(TP+FP)}$$

Presisi merupakan ukuran yang digunakan untuk menilai tingkat ketepatan model dalam mengklasifikasikan data sebagai kelas positif. Nilai presisi yang tinggi menunjukkan bahwa sebagian besar hasil prediksi positif yang diberikan model benar adanya.

3. Recall

$$Recall = \frac{(TP)}{(TP+FN)}$$

Recall menggambarkan kemampuan model dalam mengidentifikasi seluruh data yang benar-benar termasuk ke dalam kelas positif. Nilai recall yang tinggi menunjukkan bahwa model mampu menangkap sebagian besar data positif tanpa banyak yang terlewat.

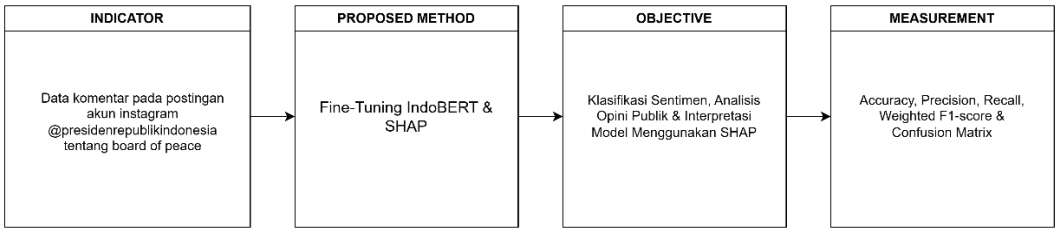
4. F1-Score

$$F1 - Score = \frac{(2 \times Recall \times Precision)}{(Recall+Precision)}$$

F1-Score adalah nilai rata-rata harmonik dari precision dan recall yang digunakan untuk menyeimbangkan keduanya. Metrik ini sangat bermanfaat ketika diperlukan keseimbangan antara ketepatan dan kelengkapan dalam hasil klasifikasi.

True Positive (TP) menunjukkan prediksi positif yang sesuai dengan kondisi sebenarnya, sedangkan *True Negative* (TN) menunjukkan prediksi negatif yang benar. *False Positive* (FP) adalah kesalahan Tipe I, yang terjadi ketika data negatif dianggap positif. Sebaliknya, *False Negative* (FN) adalah kesalahan Tipe II, yang terjadi ketika data positif tidak dapat diidentifikasi suatu kondisi yang biasanya memiliki risiko yang lebih tinggi (Arvio et al., 2024).

2.3 Kerangka Pemikiran



Gambar 2. 4 Tahapan Kerangka Penelitian

Gambar di atas menggambarkan kerangka kerja penelitian dengan alur metodologi yang sistematis. Penjelasan tiap tahap adalah sebagai berikut:

1. *Indicator*

Data yang digunakan dalam penelitian ini berupa komentar pengguna pada postingan Instagram akun @presidenrepublikindonesia yang membahas partisipasi Indonesia dalam *Board of Peace*. Data tersebut diperoleh melalui proses scraping komentar media sosial, sehingga mencerminkan opini publik secara langsung terkait isu yang diteliti.

2. *Proposed Method*

Penelitian ini mengusulkan penerapan *Fine-tuning IndoBERT* sebagai metode utama untuk mengklasifikasikan sentimen komentar ke dalam tiga kategori, yaitu positif, negatif, dan netral. Selanjutnya, metode SHAP (*SHapley Additive exPlanations*) diterapkan untuk menginterpretasikan hasil prediksi model dengan menunjukkan kontribusi setiap token terhadap keputusan klasifikasi, baik secara global maupun lokal.

3. *Objective*

Penelitian ini bertujuan mengklasifikasikan sentimen komentar Instagram menggunakan model *IndoBERT* untuk memperoleh gambaran opini publik terhadap kebijakan *Board of Peace*. Selain itu, penelitian ini bertujuan menginterpretasikan hasil prediksi model menggunakan metode SHAP sehingga proses pengambilan keputusan model menjadi lebih transparan dan mudah dipahami.

4. *Measurement*

Tahap *measurement* dilakukan untuk mengukur kinerja model menggunakan *Accuracy, Precision, Recall, Weighted F1-score,* dan *Confusion matrix*. Selanjutnya, metode SHAP digunakan untuk menginterpretasikan hasil prediksi melalui analisis global dan lokal, sedangkan *Word Cloud* digunakan sebagai visualisasi pendukung untuk menggambarkan kata-kata yang dominan pada setiap kelas sentimen.