BAB 2
LANDASAN TEORI

2.1 Analisis Sentimen pada Bahasa Indonesia

Analisis sentimen merupakan cabang dari pemrosesan bahasa alami (Natural Language Processing) yang bertujuan untuk mengekstraksi opini, emosi, dan evaluasi subjek terhadap suatu topik dari data tekstual [8]. Dalam konteks Bahasa Indonesia, analisis sentimen menghadapi tantangan unik akibat spektrum formalitas yang luas, mulai dari teks berita formal hingga penggunaan bahasa kolokial, slang, dan singkatan yang dinamis di media sosial [8]. Penggunaan model berbasis Deep Learning tingkat lanjut diperlukan untuk menangkap makna kontekstual dan nuansa bahasa tersirat yang seringkali gagal dikenali oleh metode berbasis leksikon atau pembelajaran mesin klasik [1].

2.2 Arsitektur Transformer dan BERT

Bidirectional Encoder Representations from Transformers (BERT) adalah model bahasa yang memanfaatkan arsitektur Transformer berbasis mekanisme *self-attention* [9]. Berbeda dengan model sekuensial tradisional yang memproses teks dari kiri ke kanan, BERT dilatih secara bidirectional, yang memungkinkannya mempelajari representasi kata berdasarkan konteks kiri dan kanannya secara simultan pada seluruh lapisan [10]. BERT menggunakan dua tujuan utama dalam fase pre-training: Masked Language Modeling (MLM) untuk memahami struktur kalimat melalui prediksi token yang disembunyikan, dan Next Sentence Prediction (NSP) untuk memahami hubungan antar-kalimat [11].

2.3 IndoBERT

IndoBERT adalah model bahasa berbasis arsitektur BERT yang dioptimalkan secara khusus untuk karakteristik Bahasa Indonesia [12]. Model ini dilatih menggunakan korpus Indo4B yang sangat besar, terdiri dari sekitar 4 miliar kata (kurang lebih 23 GB teks bersih) yang dikumpulkan dari berbagai sumber seperti Wikipedia Indonesia, artikel berita, dan teks media sosial [13]. Terdapat beberapa varian utama model ini, termasuk IndoBERT-base yang memiliki 12 lapisan transformer dengan 124,5 juta parameter, serta varian IndoBERT-lite yang

menggunakan arsitektur ALBERT untuk efisiensi memori melalui teknik parameter sharing [13]. IndoBERT telah terbukti memiliki performa state-of-the-art pada berbagai tugas pemahaman bahasa di Indonesia, termasuk klasifikasi sentimen [12]. Meskipun begitu, kinerja IndoBERT dalam tugas klasifikasi sentimen sangat bergantung pada kualitas dan kuantitas data latih yang digunakan [1]. Ketika data latih yang tersedia terbatas atau mengandung label yang tidak akurat, model ini dapat mengalami penurunan performa yang signifikan, sehingga diperlukan strategi peningkatan kualitas data seperti otomatisasi pelabelan (automated labelling) untuk mengoptimalkan potensi model [14].

2.4 Mekanisme Leksikon

Metode leksikon atau *lexicon-based* merupakan pendekatan analisis sentimen yang mengandalkan kamus atau daftar kata (*lexicon*) yang telah diberi bobot sentimen untuk mengklasifikasikan opini dalam suatu teks [15]. Metode ini tidak memerlukan proses pelatihan model karena hanya mencocokkan kata-kata dalam teks dengan kamus sentimen yang tersedia.

Salah satu kamus sentimen bahasa Indonesia yang paling banyak digunakan adalah InSet Lexicon (Indonesian Sentiment Lexicon) yang dikembangkan oleh Koto dan Rahmaningtyas (2017) [16]. InSet Lexicon dibangun dari kumpulan kata-kata yang diambil dari tweet berbahasa Indonesia, dengan setiap kata diberi bobot sentimen secara manual dan diperkaya dengan penambahan *stemming* serta sinonim. Kamus ini dirancang khusus untuk mengidentifikasi opini tertulis dan mengategorikannya ke dalam opini positif atau negatif [16].

Cara kerja dari metode leksikon adalah setiap kata dalam teks yang telah di *preprocessing* dicocokkan dengan kamus sentimen yang tersedia (seperti InSet Lexicon) untuk mendapatkan skor sentimen per katanya. Setelah itu, skor sentimen akan dijumlahkan untuk menentukan keseluruhan sentimen dari teks tersebut dan berdasarkan total skor yang diperoleh, teks diklasifikasikan ke dalam kategori sentimen tertentu (positif, netral, dan negatif) [16].

2.5 Mekanisme Pseudo-Labeling

Pseudo-labelling adalah teknik dalam pembelajaran *semi-supervised learning* di mana model yang telah dilatih pada data berlabel digunakan untuk memberikan label otomatis pada data baru yang belum berlabel [5]. Label yang

dihasilkan oleh model ini disebut sebagai *pseudo-label*. Pendekatan ini muncul sebagai solusi atas kendala utama dalam analisis sentimen, yaitu pelabelan manual yang tidak efisien dan memakan waktu, terutama untuk data yang sangat besar [5].

Penerapan *pseudo-labelling* dalam penelitian ini dimulai dengan pemilihan model `mdhugol/indonesia-bert-sentiment-classification` sebagai *teacher* model (model yang dipakai untuk menghasilkan label). Model yang telah melalui proses *fine-tuning* tersebut bertugas untuk klasifikasi sentimen bahasa indonesia ini digunakan untuk melakukan inferensi terhadap seluruh komentar Youtube yang belum memiliki label, sehingga menghasilkan tiga kategori sentimen (positif, netral, dan negatif) [3]. Komentar Youtube yang sudah terlabel kemudian digunakan menjadi data latih untuk melatih model IndoBERT *Base Uncased* yang berperan sebagai *student* model [3]. Dengan proses ini, model akhir dapat memperoleh pemahaman domain yang lebih baik tanpa harus melalui proses pelabelan manual yang memakan waktu dan sumber daya.

Pendekatan *pseudo-labelling* menawarkan kelebihan utama berupa efisiensi waktu dan skalabilitas, karena teknik ini mampu melabeli ribuan sampel data dalam waktu yang relatif singkat serta memungkinkan pemanfaatan data tidak berlabel yang jumlahnya jauh lebih besar dan lebih mudah diperoleh dibandingkan data berlabel manual [3]. Namun, penerapan teknik ini juga menghadapi sejumlah tantangan yang perlu diantisipasi. Kualitas pseudo-label yang dihasilkan sangat bergantung pada performa teacher model yang digunakan; jika teacher model mengandung bias atau kelemahan tertentu, maka bias tersebut akan terbawa ke dalam data latih dan berpotensi menurunkan kualitas model akhir [3]. Oleh karena itu, pemilihan model *teacher* merupakan hal penting dalam implementasi teknik *pseudo-labelling*

2.6 Penanganan Data Tidak Seimbang (Class Weighting)

Data yang didapatkan dari terkumpulnya komentar Youtube menunjukkan ketidakseimbangan antar kelas sentimen. Ketidakseimbangan kelas tersebut sering ditemukan pada dataset media sosial, di mana jumlah sampel pada satu kategori sentimen jauh melampaui kategori lainnya [17]. Tanpa penanganan, model cenderung mengalami bias terhadap kelas mayoritas dan gagal mengenali pola pada kelas minoritas. Teknik *Inverse-Frequency Class Weighting* diimplementasikan dengan menghitung bobot untuk setiap kelas dan menyuntikkannya ke dalam fungsi `CrossEntropyLoss`. Pendekatan ini secara matematis memberikan penalti *loss*

yang lebih besar jika model salah mengklasifikasikan kelas minoritas, sehingga memaksa model untuk meningkatkan sensitivitas (*recall*) terhadap kategori sentimen yang jumlahnya lebih sedikit [10]. Oleh karena itu, penerapan *class weighting* disertai *pseudo-labelling* memastikan bahwa model akhir mendapatkan hasil yang akurat dan seimbang dalam menangani seluruh kategori kelas sentimen.