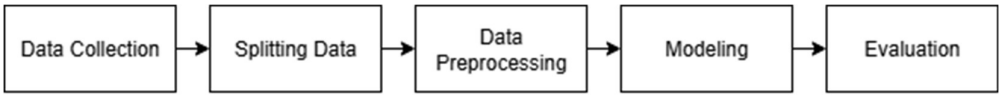
BAB III

METODE PENELITIAN

3.1. Metode Penelitian

Metode utama yang diusulkan dalam penelitian ini adalah klasifikasi emosi menggunakan model *IndoBERTweet* yang dioptimalkan dengan *Focal Loss*. Untuk mengimplementasikan dan menguji metode ini, penelitian disusun melalui beberapa tahapan sistematis yang berfokus pada data Twitter berbahasa Indonesia. *Focal Loss* diterapkan sebagai loss function yang dimodifikasi untuk mengatasi tantangan ketidakseimbangan kelas emosi, dengan tujuan meningkatkan performa model, terutama pada kelas-kelas emosi minoritas. Alur kerja penelitian, mulai dari pengumpulan data hingga evaluasi, dibangun berdasarkan tahapan-tahapan standar dalam penelitian klasifikasi teks berbasis transformer. Rangkaian tahapan penelitian yang dilakukan diuraikan secara rinci pada bagian-bagian berikut, sebagaimana diilustrasikan pada Gambar 3.1.



Gambar 3. 1 Tahapan Penelitian

3.1.1. Pengumpulan Data

Penelitian ini menggunakan data sekunder berupa dataset publik yang telah tersedia. Dataset yang dipilih adalah "Emotion dataset from Indonesian public opinion" (Riccosan dkk., 2022). Dataset ini dipilih karena beberapa alasan strategis: pertama, sangat relevan dengan topik penelitian, yaitu klasifikasi emosi dari data

Twitter berbahasa Indonesia. Kedua, dataset ini telah digunakan sebagai tolok ukur (*benchmark*) oleh beberapa penelitian sebelumnya, sehingga memungkinkan adanya perbandingan kinerja model secara langsung dan adil.

Menurut deskripsi dari penulisnya, dataset ini dibangun dengan mengumpulkan data dari platform media sosial Twitter menggunakan Twitter API. Pengumpulan data asli difokuskan pada tweet berbahasa Indonesia (lang='id') dengan menggunakan kata kunci spesifik yang berasosiasi dengan emosi tertentu. Setelah melalui proses pembersihan dan anotasi manual oleh tiga orang, dataset final berisi 7.080 tweet yang telah diberi label ke dalam enam kategori emosi, yaitu *anger* (marah), *fear* (takut), *joy* (senang), *love* (cinta), *sadness* (sedih), dan *neutral* (netral). Dataset ini tersedia secara publik dan dapat diakses melalui repositori GitHub, yang mendukung prinsip reproduktifitas dalam penelitian.

3.1.2. Pembagian Data

Dataset pada penelitian ini dibagi menjadi dua bagian, yaitu data latih dan data uji, dengan menggunakan metode *stratified split* agar distribusi kelas emosi tetap seimbang. Pembagian data ini dilakukan untuk mencegah terjadinya *data leakage*, yaitu kondisi ketika informasi dari data uji ikut memengaruhi proses pelatihan model. Jika hal tersebut terjadi, maka performa model dapat terlihat tinggi secara semu karena model sudah mengenali data yang seharusnya baru ditemui pada tahap pengujian (Bushuiev dkk., 2024). Selain itu, pembagian data juga penting untuk mengurangi risiko *overfitting*, yaitu ketika model terlalu menyesuaikan diri dengan pola pada data latih sehingga gagal melakukan generalisasi pada data baru (Joeres dkk., 2023). Dengan adanya data uji yang benar-

benar terpisah dari proses pelatihan, evaluasi performa dapat dilakukan secara lebih objektif. Pada penelitian ini, dataset dibagi dengan proporsi 70% data latih, 15% data validasi dan 15% data uji**.**

3.1.3. Pra-pemrosesan Data

Tahap pra-pemrosesan data dilakukan untuk menyiapkan data teks agar dapat diproses oleh model IndoBERTweet. Dataset yang digunakan pada penelitian ini merupakan *Emotion Dataset from Indonesian Public Opinion* yang dikembangkan oleh (Riccosan dkk., 2022). Dataset tersebut telah melalui proses pembersihan data (*data cleaning*) pada tahap pembentukannya, yang meliputi penghapusan data duplikat, pengubahan huruf menjadi bentuk kecil (*lowercasing*), serta penghapusan elemen yang tidak relevan seperti *hashtag*, *mention*, tautan (*URL*), emotikon, tanda baca, simbol non-emosional (misalnya "@", "%", "$"), dan karakter berlebih seperti spasi atau tanda baca ganda.

Proses pembersihan tersebut telah dilakukan oleh pengembang dataset, penelitian ini tidak melakukan pembersihan tambahan agar struktur dan konteks asli tweet tetap terjaga. Keputusan ini didasarkan pada karakteristik model IndoBERTweet yang secara khusus dilatih menggunakan korpus Twitter berbahasa Indonesia, yang mencakup gaya bahasa informal dan penggunaan emotikon khas media sosial (Koto dkk., 2021). Dengan demikian, penghapusan unsur informal justru berpotensi menghilangkan informasi kontekstual penting yang berkaitan dengan ekspresi emosi.

Tahap pra-pemrosesan dalam penelitian ini difokuskan pada proses *tokenisasi* dan penyiapan format input agar sesuai dengan kebutuhan arsitektur *IndoBERTweet*. Proses tokenisasi dilakukan menggunakan *IndoBERTweet tokenizer* untuk mengubah setiap teks menjadi token-token numerik berupa *input IDs* dan *attention mask*. Tokenisasi ini juga menambahkan token khusus `[CLS]` di awal dan `[SEP]` di akhir kalimat, sesuai dengan standar struktur input pada model berbasis Transformer.

Selanjutnya dilakukan proses *padding* dan *truncation* untuk menyeragamkan panjang setiap urutan token menjadi maksimum 128 token, berdasarkan rata-rata panjang teks pada korpus pelatihan *IndoBERTweet*. Proses ini memastikan setiap input memiliki dimensi yang konsisten agar dapat diproses secara batch pada tahap pelatihan model.

Selain pra-pemrosesan terhadap teks, dilakukan pula konversi terhadap label emosi agar dapat dikenali oleh model klasifikasi. Setiap label emosi yang semula berbentuk teks, yaitu *marah, takut, senang, cinta, sedih,* dan *netral*, diubah menjadi representasi numerik yang unik menggunakan teknik *label encoding*. Proses ini bertujuan untuk memudahkan model dalam membedakan antar kelas emosi secara numerik selama pelatihan dan evaluasi. Hasil akhir dari tahap pra-pemrosesan ini berupa pasangan data (*input IDs*, *attention mask*) beserta label emosi numerik yang siap digunakan pada tahap pelatihan model *IndoBERTweet*.

3.1.4. Pemodelan

Pemodelan dalam penelitian ini menggunakan pendekatan *transformer-based model* dengan *IndoBERTweet* sebagai model dasar. Pemodelan dibagi ke dalam empat bagian, yaitu arsitektur model, fungsi loss yang digunakan, skenario eksperimen, serta optimizer dan proses pelatihan.

3.1.4.1.Arsitektur Model

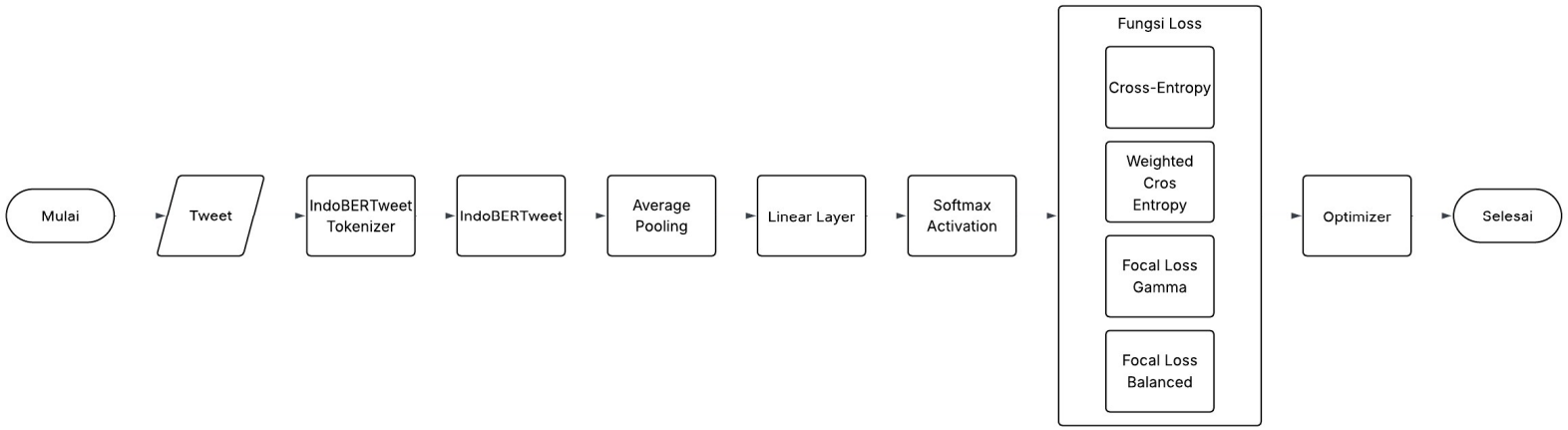
Model yang digunakan dalam penelitian ini adalah IndoBERTweet, yaitu model *language representation* berbasis arsitektur Transformer yang secara khusus dilatih menggunakan korpus Twitter berbahasa Indonesia. Model ini dipilih karena kemampuannya dalam memahami gaya bahasa informal, singkatan, serta ekspresi khas media sosial yang umum ditemukan pada data tweet (Koto dkk., 2021). Dalam penelitian ini, IndoBERTweet digunakan sebagai *feature extractor* untuk menghasilkan representasi kontekstual dari teks masukan yang kemudian digunakan untuk proses klasifikasi emosi.

Setiap teks masukan terlebih dahulu diproses menggunakan *IndoBERTweet tokenizer* untuk diubah menjadi token numerik yang sesuai dengan kosakata (*vocabulary*) model. Proses tokenisasi ini juga menambahkan token khusus `[CLS]` di awal dan `[SEP]` di akhir teks untuk menandai batas kalimat serta menyiapkan struktur input yang sesuai dengan arsitektur model (Devlin dkk., 2019). Token `[CLS]` digunakan sebagai penanda representasi keseluruhan teks, sedangkan `[SEP]` berfungsi sebagai pemisah antar kalimat atau penanda akhir teks.

Berbeda dengan pendekatan umum pada BERT yang hanya menggunakan representasi token [CLS], IndoBERTweet menerapkan strategi average pooling untuk memperoleh representasi teks secara keseluruhan. Pada tahap ini, seluruh

vektor keluaran (*hidden states*) dari lapisan terakhir Transformer dirata-ratakan sehingga menghasilkan representasi global yang lebih stabil dan informatif terhadap keseluruhan konteks kalimat (Koto dkk., 2021). Representasi hasil *average pooling* inilah yang digunakan sebagai masukan ke lapisan klasifikasi linear untuk menentukan kategori emosi yang sesuai.

Struktur umum alur pemodelan yang digunakan dalam penelitian ini ditunjukkan pada Gambar 3.2. Gambar tersebut menggambarkan aliran proses dari teks tweet hingga keluaran prediksi emosi, yang terdiri dari tahap tokenisasi, pemrosesan oleh IndoBERTweet, pengambilan representasi `[CLS]`, proses klasifikasi dengan lapisan linear, serta pelatihan menggunakan *Focal Loss* dan algoritma *optimizer* Adam.



Gambar 3. 2 Aliran Proses Pemodelan

Seperti yang ditunjukkan pada Gambar 3.2, hasil keluaran dari token `[CLS]` digunakan sebagai representasi global teks yang kemudian diteruskan ke lapisan klasifikasi linear (*fully connected layer*) dengan enam neuron keluaran. Setiap neuron mewakili satu kelas emosi, yaitu *marah, takut, senang, cinta, sedih,* dan *netral*. Sebelum lapisan klasifikasi akhir, diterapkan lapisan *dropout* dengan nilai 0.3 untuk mengurangi risiko *overfitting* dan meningkatkan kemampuan generalisasi model terhadap data uji. Lapisan linear ini berfungsi memetakan vektor berdimensi 768 menjadi enam nilai logit, yang kemudian diubah menjadi probabilitas melalui fungsi

aktivasi *softmax* (Kim, 2016). Kelas dengan probabilitas tertinggi ditetapkan sebagai hasil prediksi akhir model.

Selama proses pelatihan, digunakan fungsi *Focal Loss* untuk menghitung perbedaan antara hasil prediksi dan label aktual. Fungsi ini merupakan pengembangan dari *Cross-Entropy Loss* yang menambahkan dua parameter utama, yaitu faktor alpha (α) sebagai pembobot kelas dan faktor gamma (γ) sebagai pengontrol tingkat fokus pada data yang sulit diklasifikasikan (Lin dkk., 2018). Dengan demikian, *Focal Loss* membantu mengatasi ketidakseimbangan kelas (*class imbalance*) pada dataset emosi. Nilai *loss* yang dihasilkan kemudian digunakan oleh algoritma *optimizer* Adam untuk memperbarui bobot model secara iteratif hingga mencapai konvergensi (Loshchilov & Hutter, 2019).

3.1.4.2.*Focal Loss*

Fungsi *Focal Loss* digunakan dalam penelitian ini untuk mengatasi ketidakseimbangan kelas pada data emosi. Fungsi ini merupakan pengembangan dari *Cross-Entropy Loss* yang menambahkan dua parameter utama, yaitu faktor α sebagai pembobot antar kelas dan faktor γ sebagai pengontrol fokus terhadap sampel yang sulit diklasifikasikan (*Lin, Goyal, Girshick, He, & Dollár, 2017*). Nilai γ berfungsi untuk mengurangi pengaruh sampel yang mudah diklasifikasikan, sehingga model lebih menekankan pembelajaran pada sampel yang sulit atau kelas minoritas.

Dalam penelitian *Focal Loss* yang diterapkan pada klasifikasi biner, α ditetapkan sebesar 0,25 untuk kelas positif dan $1 - \alpha = 0{,}75$ untuk kelas negatif guna menyeimbangkan jumlah contoh antar kelas (*Lin et al., 2017*). Namun, karena

penelitian ini merupakan klasifikasi multi-kelas dengan enam kategori emosi yang memiliki distribusi data tidak seimbang, pendekatan tersebut tidak lagi sesuai.

Oleh karena itu, penelitian ini menggunakan pendekatan *balanced weighting* berdasarkan prinsip *inverse class frequency* sebagaimana disarankan dalam penelitian *Focal Loss* (*Lin et al., 2017*), yang menyatakan bahwa dalam praktiknya nilai α dapat ditetapkan berdasarkan *inverse class frequency* atau disetel melalui *cross-validation*. Dengan pendekatan ini, kelas yang memiliki jumlah data lebih sedikit akan mendapatkan nilai α yang lebih besar secara proporsional, sedangkan kelas dengan data yang lebih banyak akan memiliki α yang lebih kecil. Pendekatan ini memberikan kontribusi yang seimbang dalam proses perhitungan *loss* antar kelas. Formula perhitungan α yang digunakan dalam penelitian ini dituliskan sebagai berikut:

$$\boldsymbol{w_c = \frac{N}{n_c C}}$$

Persamaan 3. 1 pembobotan kelas

Dimana:

$\boldsymbol{w_c}$ = bobot kelas c

$\boldsymbol{N}$ = jumlah total sampel

$\boldsymbol{C}$ = jumlah kelas

$\boldsymbol{n_c}$ = jumlah sampel pada kelas C

Pendekatan ini membuat proses pembelajaran model menjadi lebih seimbang terhadap distribusi data yang tidak merata. Selain itu, parameter γ digunakan untuk mengatur tingkat fokus model terhadap sampel yang sulit diklasifikasikan. Penelitian ini menggunakan dua nilai γ, yaitu 0 dan 2, untuk membandingkan pengaruh tingkat fokus terhadap hasil klasifikasi emosi.

3.1.4.3.Skenario eksperimen

Pada tahap ini dilakukan serangkaian eksperimen untuk menganalisis pengaruh kombinasi parameter α (pembobot kelas) dan γ (faktor fokus) pada fungsi *loss* terhadap performa model klasifikasi emosi. Tujuan utama eksperimen ini adalah untuk menentukan konfigurasi fungsi *loss* yang paling optimal dalam menghadapi distribusi data yang tidak seimbang.

Penelitian ini merancang empat skenario eksperimen untuk membandingkan kinerja berbagai konfigurasi fungsi *loss*. Skenario pertama menggunakan *Cross-Entropy Loss* standar dengan nilai $\alpha = 1$ dan $\gamma = 0$, yang dijadikan sebagai baseline karena merupakan fungsi *loss* paling umum dan dasar dalam klasifikasi multi-kelas.

Skenario kedua menggunakan *Weighted Cross-Entropy Loss* dengan bobot kelas yang ditentukan berdasarkan prinsip *inverse class frequency*. Skenario ini dipilih untuk mengatasi ketidakseimbangan kelas dengan memberikan perhatian lebih pada kelas yang memiliki jumlah sampel lebih sedikit.

Skenario ketiga menerapkan *Focal Loss* dengan parameter $\alpha = 1$ dan $\gamma = 2$. Konfigurasi ini dirancang untuk menekankan pembelajaran pada sampel yang sulit diklasifikasikan (*Lin, Goyal, Girshick, He, & Dollár, 2017*).

Skenario keempat menggabungkan pembobotan kelas berdasarkan *inverse class frequency* dengan penekanan pada sampel sulit melalui nilai $\gamma = 2$. Skenario ini bertujuan untuk memanfaatkan keunggulan keduanya, yaitu mengurangi bias akibat distribusi data yang tidak seimbang sekaligus meningkatkan sensitivitas model terhadap sampel yang lebih sulit diprediksi.

Tabel 3. 1 Skenario eksperimen

Skenario	Fungsi Loss	α (alpha)	γ (gamma)	Keterangan
1	Cross-Entropy Loss	1	0	Baseline klasifikasi multi-kelas standar
2	Weighted Cross-Entropy Loss	Inverse class frequency	0	Mengatasi ketidakseimbangan kelas
3	*Focal Loss Gamma*	1	2	Menekankan sampel sulit diprediksi (Lin dkk., 2018)
4	*Focal Loss Balanced*	Inverse class frequency	2	Kombinasi pembobotan kelas & penekanan sampel sulit

Evaluasi keempat skenario dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score untuk memperoleh gambaran menyeluruh mengenai performa model pada setiap konfigurasi fungsi *loss*. Keempat skenario ini memberikan dasar perbandingan untuk menganalisis sejauh mana penerapan pembobotan kelas dan faktor fokus dapat meningkatkan hasil klasifikasi emosi.

3.1.4.4.Optimizer dan proses pelatihan

Proses pelatihan model dilakukan untuk menyesuaikan parameter internal jaringan terhadap data pelatihan sehingga model mampu melakukan klasifikasi emosi secara optimal. Optimizer yang digunakan dalam penelitian ini adalah *Adam (Adaptive Moment Estimation)* karena memiliki kemampuan adaptif dalam menyesuaikan laju pembelajaran untuk setiap parameter, serta dikenal stabil dan efisien dalam meminimalkan fungsi *loss* pada model berbasis *Transformer* (Dereich & Jentzen, 2024). Optimizer ini dipilih karena dapat mempercepat proses konvergensi dan mengurangi kebutuhan penyetelan *learning rate* yang kompleks.

Pelatihan dilakukan selama maksimum 50 *epoch* dengan penggunaan *early stopping* untuk mencegah *overfitting*. Mekanisme *early stopping* diterapkan dengan *patience* sebesar 3, di mana pelatihan akan dihentikan jika nilai *F1-score* pada data validasi tidak menunjukkan peningkatan selama tiga *epoch* berturut-turut. Penggunaan metrik *F1-score* sebagai acuan penghentian dini dipilih karena mampu merepresentasikan keseimbangan antara presisi dan recall, terutama pada dataset yang memiliki distribusi kelas tidak seimbang.

Model terbaik ditentukan berdasarkan nilai *F1-score* tertinggi pada data validasi. Setelah proses pelatihan selesai, model dengan performa terbaik disimpan untuk digunakan pada tahap pengujian.

3.1.5. Evaluasi

Evaluasi model dalam penelitian ini dilakukan pada data uji dengan menggunakan metrik akurasi, presisi, recall, dan F1-score yang diperoleh dari *confusion matrix*. Metrik akurasi digunakan untuk memberikan gambaran umum tingkat ketepatan prediksi model secara keseluruhan. Presisi dipakai untuk mengukur seberapa tepat model dalam memprediksi suatu kelas tertentu, sedangkan recall menilai sejauh mana model mampu mengenali seluruh sampel yang benar dalam suatu kelas. F1-score digunakan sebagai ukuran keseimbangan antara presisi dan recall, sehingga lebih representatif pada kondisi data yang tidak seimbang.

Pemilihan keempat metrik ini dimaksudkan agar evaluasi tidak hanya bergantung pada akurasi semata, tetapi juga mempertimbangkan performa model pada tiap kelas emosi. Dengan demikian, hasil evaluasi dapat memberikan gambaran yang lebih menyeluruh mengenai kekuatan dan kelemahan model dalam klasifikasi emosi berbahasa Indonesia.