BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

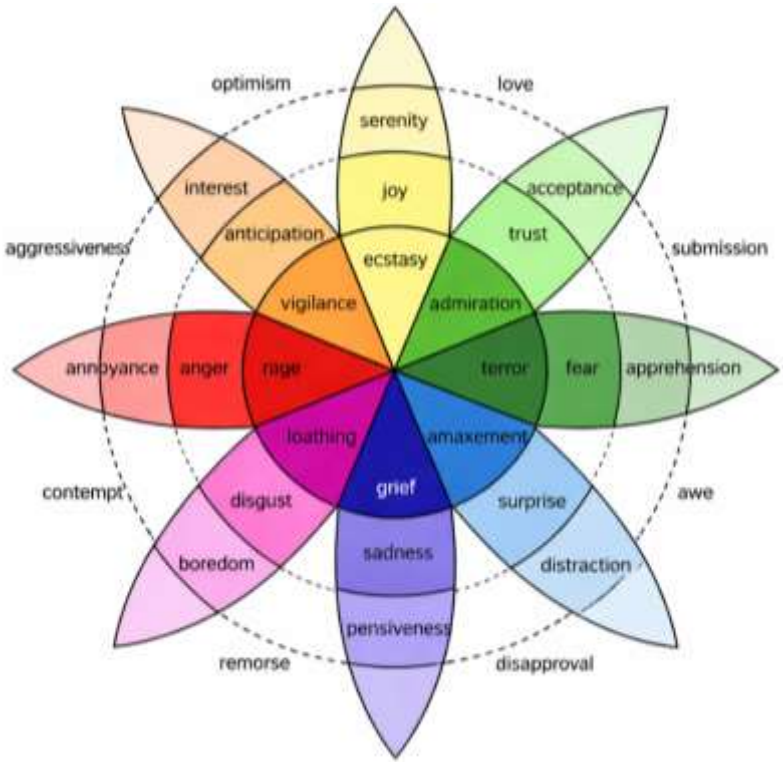
2.1.1 Emosi dalam Komunikasi dan Teks

Emosi adalah kerangka kerja bagi proses berpikir dan berfungsi sebagai landasan paling mendasar bagi proses-proses selanjutnya, yaitu fungsi tingkat tinggi dan pemrosesan emosi (Yusainy dkk., 2025). Emosi juga mencakup reaksi kompleks terhadap rangsangan internal dan eksternal, yang melibatkan perubahan fisiologis, psikologis, dan bahkan perilaku, serta sangat penting bagi komunikasi dan pengambilan keputusan manusia (Arrayan, 2025). Teori emosi dasar mencakup kategori utama universal seperti kemarahan, ketakutan, kegembiraan, kesedihan, jijik, dan keterkejutan (Pratama & Findawati, 2024) dan membentuk landasan penting bagi pengembangan model klasifikasi emosi, khususnya dalam konteks *multilingual* (Lazuardi & Noor Fatyanosa, 2025).

Emosi dapat diidentifikasi menggunakan berbagai metode, seperti suara, ekspresi wajah, dan teks (Labib dkk., 2025). Petunjuk-petunjuk ini meliputi sinyal vokal atau verbal, perilaku nonverbal, dan ekspresi wajah. Petunjuk verbal terlihat jelas dalam pemilihan kata dan cara berbicara. Contoh komunikasi verbal adalah teks. Untuk menentukan isi emosional dari teks, metode klasifikasi diterapkan (Aripin dkk., 2023). Memahami teori emosi ini sangat penting untuk meningkatkan akurasi klasifikasi emosi dalam pemrosesan bahasa alami.

2.1.2 Klasifikasi Emosi

Klasifikasi emosi mengacu pada pengelompokan emosi dan upaya untuk membedakan satu sama lain. Hal ini mencakup tugas mengenali emosi seseorang dan mengklasifikasikannya berdasarkan reaksi dan respons (Raharko & Yamasari, 2024). Klasifikasi emosi sering kali dibagi menjadi beberapa kategori utama, seperti emosi primer, emosi yang terkait dengan rangsangan sensorik, penilaian diri, dan hubungan interpersonal (Nurjannah & Suryaningsih, 2025). Pendekatan ini penting sebagai landasan untuk mengembangkan model berbasis teks *multilingual* untuk klasifikasi emosi. Teknologi *deep learning*, khususnya model berbasis *transformer*, digunakan untuk meningkatkan akurasi dan efisiensi klasifikasi emosi di berbagai konteks linguistik.



Gambar 2. 1 Plutchik wheel of emotions (Abbasi & Beltiukov, 2019)

Gambar 2.1 menggambarkan hubungan antar emosi. Setiap emosi dapat dibagi menjadi subkategori misalnya, tergantung pada konteksnya, "kebahagiaan" dapat menggambarkan keadaan kegembiraan atau kesenangan. Sebagaimana ditunjukkan oleh penelitian sebelumnya yang menggunakan *Pluctchik wheet of emotion* sebagai referensi, kategori emosi tambahan harus ditambahkan untuk merepresentasikan emosi pada *Pluctchik wheet of emotion* dengan lebih baik dan memungkinkan identifikasi emosi yang lebih tepat (Alan Tusa Bagus W, 2022).

2.1.3 Sustainable Development Goals (SDGs)

Sustainable Development Goals (SDGs) terdiri dari 17 tujuan universal yang telah disepakati oleh negara-negara di seluruh dunia untuk menciptakan kehidupan yang lebih baik dan berkelanjutan di Bumi (Ragadhita dkk., 2025). Penelitian ini menggarisbawahi relevansi SDGs, karena klasifikasi emosi multibahasa dapat mendukung pencapaian SDG 3 (Kehidupan Sehat dan Sejahtera) melalui pemantauan kesehatan mental, SDG 4 (Pendidikan Berkualitas) melalui penggunaan NLP untuk pembelajaran *multilingual*, dan SDG 5 (Kesetaraan *Gender*) melalui analisis emosi yang lebih inklusif dan tidak bias. Selain itu, penelitian ini selaras dengan tujuan nasional pemerintah (Asta Cita), khususnya poin keempat, yang bertujuan untuk memperkuat pengembangan sumber daya manusia, *sains*, teknologi, pendidikan, dan kesetaraan gender. Pengembangan model klasifikasi emosi *multilingual* berbasis kecerdasan buatan dapat berkontribusi pada peningkatan kualitas

pembelajaran, komunikasi digital, dan penggunaan teknologi yang inklusif di masyarakat (Tlogopayung, 2024).

2.1.4 Klasifikasi Teks *Multilingual*

Klasifikasi teks yaitu proses di mana komputer mengelompokkan teks ke dalam kelompok-kelompok yang terorganisir. Ini merupakan subbidang dari NLP dan dapat secara otomatis menganalisis teks serta memberikan label atau kategori yang telah ditentukan sebelumnya berdasarkan isinya. Meskipun penelitian sebelumnya mengenai klasifikasi teks sebagian besar berfokus pada teks *monolingual*, saat ini semakin banyak studi yang membahas klasifikasi teks *multilingual*. *Multilingual* mengacu pada kemampuan untuk memahami, menggunakan, dan memproses berbagai bahasa secara bersamaan dalam suatu sistem atau konteks komunikasi. Pendekatan multibahasa sangat penting untuk menangani keragaman bahasa yang digunakan dalam data teks (Nazir dkk., 2025).

2.1.5 Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan subbidang kecerdasan buatan yang berfokus pada pemrosesan bahasa alami. NLP digunakan untuk mengekstrak fitur dari teks dan menciptakan representasi semantik (Retnowati dkk., 2025). Model berbasis *deep learning* mampu memahami konteks dan nuansa emosional dalam teks, sehingga meningkatkan akurasi klasifikasi emosi dibandingkan dengan metode lain. *Natural Language Processing* (NLP) merupakan kemajuan dalam *Artificial Intelligence* (AI) yang diharapkan mampu mempelajari bahasa yang digunakan oleh manusia. NLP juga memiliki

kesempatan dalam berbagai peluang pengembangan, termasuk klasifikasi teks *multilingual*, yang menjadi subjek penelitian ini.

2.1.6 Kompleksitas Bahasa

Kompleksitas komunikasi dan kebutuhan untuk beradaptasi dengan keragaman linguistik semakin meningkat, terutama dalam konteks pemrosesan *big data* dan interaksi manusia-mesin (Nur Azizah dkk., 2025). Keragaman linguistik menimbulkan tantangan besar bagi pemrosesan *big data* karena kompleksitasnya yang tinggi, seperti morfologi aglutinatif dan variasi struktur kalimat. Kompleksitas ini mempersulit pengembangan model NLP, terutama terkait kompleksitas linguistik struktural, termasuk morfologi aglutinatif. Artinya, kata-kata dapat mengandung berbagai imbuhan yang mengubah maknanya, sehingga membuat analisis sintaksis dan semantik menjadi lebih kompleks (Ernawati dkk., 2023). Pada penelitian (Fazri & Voutama, 2025) menunjukkan bahwa panjang teks dan kompleksitas struktur kalimat secara signifikan mempengaruhi kinerja klasifikasi dalam memproses teks pendek dan padat. Pendekatan berbasis *Transformer* seperti BERT telah terbukti unggul dalam memproses teks pendek dan konteks yang kompleks (Malasari & Ramli, 2025).

2.1.7 Semantik vs Sintaksis

Semantik dan sintaksis merupakan hal mendasar dalam memahami teks multibahasa, meskipun variasi dalam pola semantik mempersulit transfer pengetahuan dalam mBERT dan XLM-RoBERTa. Semantik melibatkan pemahaman makna dalam konteks emosional. Semantik mencakup interpretasi

kata, frasa, dan kalimat, serta analisis bagaimana bahasa menyampaikan konsep, maksud, dan informasi referensial (Rakhmonova, 2025). Semantik dibagi menjadi beberapa tingkatan: semantik leksikal (makna kata-kata individual), semantik komposisional (bagaimana makna digabungkan), dan semantik formal (representasi logis dari makna) (Rakhmonova, 2025). Anotasi semantik menambahkan lapisan makna pada data yang tidak terstruktur dan mempermudah mesin untuk memahami konteks tekstual (Bhavik, 2025).

Tidak seperti sintaksis, semantik terutama memeriksa bagaimana kata-kata digabungkan menjadi kalimat yang benar secara tata bahasa. Hubungan antara sintaksis dan semantik adalah salah satu topik yang paling kompleks dan kontroversial dalam linguistik. Meskipun sintaksis mengatur bentuk dan semantik mengatur makna, bahasa alami berfungsi melalui interaksi dinamis antara kedua sistem ini. Antarmuka ini berfokus pada bagaimana konfigurasi struktural bahasa berkorelasi dengan makna (Rakhmonova, 2025).

2.1.8 Pra-Pemrosesan Data

Pra-pemrosesan data melibatkan persiapan dan transformasi data mentah menjadi format yang lebih terstruktur dan sesuai untuk analisis (Ishak, 2023). Data yang dikumpulkan melalui tahap pra-pemrosesan yang mencakup normalisasi, pembersihan data, dan ekstraksi fitur (Kamal, 2025). Pra-pemrosesan data sangat penting karena membantu mengurangi variasi kata yang tidak perlu, menyederhanakan data, serta meningkatkan akurasi dan efisiensi model untuk analisis sentimen, klasifikasi teks, dan ekstraksi informasi (Wahyu dkk., 2026).

2.1.9 *Data Balancing*

Data balancing adalah teknik pemrosesan data yang dirancang untuk mengatasi ketidakseimbangan antara kelas mayoritas dan minoritas (Admin Stis, 2025). Ketidakseimbangan ini sering terjadi dalam berbagai skenario. Penyeimbangan data mengevaluasi data yang akan digunakan untuk memastikan keseimbangannya (Akbar & Hayaty, 2020). Salah satu pendekatan yang digunakan dalam menangani data tidak seimbang adalah *cross-entropy loss* dengan *class weight loss,* yaitu metode yang memberikan bobot berbeda pada setiap kelas dalam fungsi loss (Razali dkk., 2025). Pembobotan *class weight* dicapai dengan fungsi *loss* sehingga kelas minoritas menerima bobot yang lebih tinggi daripada kelas mayoritas. Teknik penyeimbangan data ini berfungsi untuk mengurangi bias prediksi dan meningkatkan akurasi secara keseluruhan, terutama pada klasifikasi *multiclass* yang tidak seimbang.

2.1.10 Akuisisi Data

Akuisisi data merupakan salah satu tahap dalam metodologi pengembangan data (Huber dkk., 2019). Tahap ini mencakup pengumpulan data mentah yang diperlukan untuk proyek pengembangan data. Data mentah tersebut kemudian diolah menjadi data akhir melalui berbagai tahap analisis dan pemrosesan data. Data akhir ini kemudian digunakan dalam proses pemodelan (Eko Wahyudi dkk., 2022).

2.1.11 Konversi Emoji

Emoji telah menjadi sangat populer dalam komunikasi digital. Penggunaannya dalam komunikasi daring dipandang sebagai cara bagi

pengguna untuk menambahkan sentuhan yang lebih personal pada pesan teks mereka (Mladenović dkk., 2023). Mengintegrasikan deskripsi teks untuk emoji ke dalam dataset dapat secara signifikan meningkatkan akurasi klasifikasi sentimen dan emosi. Hal ini karena model dapat mengidentifikasi indikator emosional spesifik melalui label teks yang unik, bukan sekadar memproses data mentah tanpa simbol (Reddy dkk., 2025). Konversi emoji dapat dilakukan menggunakan pustaka python.

2.1.12 Penghapusan Simbol atau Karakter

Proses *filtering* merupakan langkah penting dalam pra-pemrosesan data. Teks masukan disaring untuk mendapatkan satu set data tunggal guna diproses lebih lanjut (Sujjada & Fergina, 2021). Hal ini melibatkan penghapusan elemen yang tidak relevan seperti URL, penyebutan pengguna (*@username*), tagar, angka, dan karakter khusus (selain huruf dan spasi) (Karim & Wibowo, 2025). Langkah ini mengurangi gangguan dalam teks, sehingga sistem dapat berfokus pada inti teks.

2.1.13 Normalisasi

Normalisasi digunakan untuk mempersiapkan data guna mengoptimalkan penggunaannya sesuai kebutuhan pengguna serta mendukung proses-proses selanjutnya yang bertujuan mencapai hasil yang lebih baik (Kartika & Maulana, 2021). Tujuan normalisasi teks adalah untuk menstandarkan struktur kalimat guna memastikan konsistensi dan menyederhanakan pengelolaan data sistem. Dalam proses ini, kata-kata yang tidak baku diubah menjadi bentuk bakunya, misalnya, “gk” menjadi “tidak” atau “sy” menjadi “saya” (Wahyu dkk., 2026).

2.1.14 Tokenisasi

Tokenisasi adalah proses memecah kalimat menjadi kata. Dalam proses ini, kata-kata dipisahkan berdasarkan spasi, dan tanda baca yang tidak diperlukan dihilangkan (Handayani dkk., 2022). Tokenisasi juga dapat didefinisikan sebagai proses memecah string masukan menjadi kata-kata individualnya. Proses ini menghasilkan token yang mewakili teks asli dan berfungsi sebagai dasar untuk mengembangkan fitur teks (Sujjada & Fergina, 2021). Dalam model XLM-RoBERTa, tokenisasi dilakukan menggunakan *SentencePiece tokenizer* berbasis *byte-pair encoding* (BPE) dengan kosa kata 250.000 token yang mendukung 100 bahasa (Conneau dkk., 2020).

2.1.15 *Padding*

Padding dilakukan untuk menstandarkan panjangnya. Proses ini dilakukan dengan menggunakan *tokenizer* (Fauzul dkk., 2017) . Proses *padding* bertujuan untuk untuk memastikan bahwa semua urutan teks memiliki panjang yang sama sehingga dapat diproses secara seragam oleh model.

2.1.16 *Fine-tuning*

Fine-tuning adalah teknik yang memanfaatkan model yang telah dilatih untuk tugas tertentu, kemudian lapisan-lapisan terakhir dari model tersebut digabungkan untuk menyelesaikan tugas baru. *Fine-tuning* dapat meningkatkan kinerja model secara signifikan (Bert dan DistilBERT 2091). *Fine-tuning Multilingual* BERT (mBERT) dilakukan karena model ini pada dasarnya dilatih pada korpus teks umum, sehingga representasi bahasa yang dihasilkan tetap bersifat umum. Lapisan klasifikasi ditambahkan di bagian

akhir arsitektur mBERT untuk memprediksi label emosi dari teks masukan (Ricciardi & Manisera, 2025).

Kemampuan model klasifikasi untuk menghasilkan keluaran probabilitas dengan nilai antara 0 dan 1 ditentukan oleh *cross-entropy loss*. *Fine-tuning* mBERT pada penelitian ini dilakukan dengan mengoptimasi *cross-entropy loss*, ketika output diberikan sebagai probabilitas, fungsi *loss* ini sesuai untuk digunakan. Fungsi loss dilihat pada persamaan (2.1).

$$L = \sum_{i=1}^{k} y_i \log (o_i) \qquad (2.1)$$

Keterangan:

a. y_i adalah label dari klasifikasi
b. o_i adalah probabilitas yang diprediksi oleh model terhadap label
c. $\log(o_i)$ adalah nilai logaritma dari tiap probabilitas yang diprediksi oleh model.

Adam adalah salah satu metode optimasi yang paling banyak digunakan, estimasi momen dihitung dan digunakan untuk mengoptimalkan fungsi adam atau algoritma estimasi model adaptif (Tato dkk., 2018). Menggunakan optimizer AdamW, persamaan (2.2).

$$\theta_t = \theta_{t-1} - \alpha(\nabla L + \gamma\theta_{t-1}) \qquad (2.2)$$

Keterangan:

a. θ_t , parameter model pada itersi ke-t
b. θ_{t-1}, parameter sebelumnya

c. α, *learning rate* yaitu besar langkah pembelajaran.

d. ∇L, gradien dari fungsi loss

e. $\gamma\theta_{t-1}$, komponen regularisasi L2 (*weight decay*)

Model *multilingual* kedua yang digunakan pada penelitian ini adalah XLM-RoBERTa. Model *transformer multilingual* ini didasarkan pada arsitektur RoBERTa dan dilatih menggunakan kumpulan teks besar dalam berbagai bahasa (Ricciardi & Manisera, 2025). Proses *fine-tuning* dilakukan dengan melatih ulang model pada dataset selama beberapa *epoch* menggunakan optimizer *AdamW* sesuai pada persamaan (2.2) dengan *learning rate sebesar* 0.00002 (Almalki, 2025).

2.1.17 *Cross-lingual*

Cross-lingual adalah metode yang memanfaatkan bahasa dengan sumber daya melimpah untuk mengatasi keterbatasan dataset pada bahasa-bahasa dengan sumber daya terbatas (Conneau dkk., 2020). Studi-studi sebelumnya menunjukkan bahwa *cross-lingual transfer learning* dapat membantu mengatasi keterbatasan dataset, terutama pada bahasa-bahasa dengan sumber daya terbatas (A. Kumar & Albuquerque, 2021), (Khairunnisa dkk., 2023). Karya ini berfokus pada penerapan pembelajaran *cross-lingual* dalam klasifikasi emosi. Berbeda dengan analisis sentimen, yang umumnya mengklasifikasikan teks berdasarkan kategori seperti positif, negatif, dan netral, klasifikasi emosi melibatkan pengenalan kategori emosi utama seperti kegembiraan, kemarahan, kesedihan, ketakutan, cinta, dan sebagainya. Hal ini menjadikan klasifikasi emosi sebagai tugas klasifikasi *multi-class* yang lebih kompleks, karena emosi memiliki dimensi semantik yang seringkali ambigu

dan bergantung pada konteks budaya dan linguistik. Meskipun model *multilingual* berbasis *transformer* telah berhasil diterapkan untuk tugas-tugas NLP lainnya, penting untuk mengevaluasi kinerja dan adaptabilitasnya secara khusus untuk tugas klasifikasi emosi, terutama dalam konteks bahasa dengan sumber daya data yang terbatas.

2.1.18 *Transformer*-BERT

Model berbasis *transformer* adalah jenis arsitektur model *deep learning* yang telah menjadi sangat popular dalam *Natural Language Processing* (NLP) dan bidang lain (M.Prytula, 2024). BERT singkatan dari *Bidirectional Encoder Representations from Transformers* yaitu kerangka kerja pembelajaran mesin *open-source* yang dikembangkan oleh Google (Salve & Tubil, 2025), BERT mendukung representasi bahasa *universal multilingual* untuk 104 bahasa. Model ini dilatih sebelumnya menggunakan 2,5 miliar kata teks Wikipedia yang tidak berlabel dan 800 juta kata dari korpus buku untuk memperoleh *embedding* kontekstual (M.Prytula, 2024).

Inti dari model-model ini adalah mekanisme *self-attention*, yang memungkinkan model menilai pentingnya bagian-bagian berbeda dari urutan masukan saat memproses setiap token. Model *transformer* didasarkan pada arsitektur *encoder-decoder*. *Encoder* memproses urutan *input* untuk menciptakan representasi berdimensi tetap, sementara *decoder* menghasilkan urutan *output* (M.Prytula, 2024). Model *transformers* telah merevolusi tugas NLP dengan secara efisien mempelajari hubungan kontekstual dalam urutan, menggunakan mekanisme perhatian, dan karenanya menjadi sangat penting

bagi aplikasi kecerdasan buatan modern, model-model ini biasanya digunakan dalam dua tahap yaitu *pre-training* dan *fine-tuning* (M.Prytula, 2024).

Transformer-BERT telah terbukti sebagai model yang unggul dalam berbagai tugas NLP berkat kemampuannya menangkap konteks kata secara dua arah (Retnowati dkk., 2025). BERT dikembangkan untuk membantu komputer memahami makna linguistik yang ambigu dalam teks dengan menggunakan *transformer* untuk membangun konteks antara kata-kata dari teks di sekitarnya (Cameron Hashemi-Pour, 2024). Kemampuan BERT untuk secara bersamaan memahami konteks kiri dan kanan merupakan pergeseran paradigma dalam pengembangan model bahasa, yang memungkinkan kemajuan signifikan dalam tugas-tugas pemahaman bahasa (Zhou dkk., 2023).

Proses *pre-training* BERT melibatkan dua tugas prediksi. Tugas pertama adalah *Masked Language Model* yang menyembunyikan kata-kata tertentu selama pelatihan dan berusaha mengidentifikasi kata yang hilang. Tugas kedua adalah *Next Sentence Prediction* yang menentukan apakah kalimat kedua mengikuti kalimat pertama (M.Prytula, 2024). BERT mengkodekan pengetahuan sintaksis dan semantic untuk menentukan kemampuannya dalam memproses struktur hierarkis, yang memainkan peran kunci dalam menyelesaikan ambiguitas di berbagai kosakata (Wang dkk., 2023), (Lyu dkk., 2023), (Samuel dkk., 2023). BERT yaitu kalimat yang dikodekan menggunakan *embedding* konteks dari model BERT yang telah dilatih sebelumnya (Gundam, Marri, dan Mamidi, 2025). Input ke model BERT adalah sebuah kalimat, yang dikonversi menjadi vektor per kata dalam kalimat

tersebut, yang umumnya disebut sebagai token (Gundam, Marri, dan Mamidi, 2025). Dalam model BERT, token direpresentasikan menggunakan model *embedding* token yang berbeda di dalam *Transformer*, yaitu melalui *embedding* posisional yang mewakili posisi token dalam urutan. Token pertama dari setiap urutan adalah token klasifikasi khusus [CLS], diikuti oleh token dari teks pertama, kemudian token pemisah [SEP], setelah itu token teks kedua, dan diulang sampai selesai (Basbeth & Fudholi, 2024).

Model BERT terdiri dari dua tahap utama *pre-training* dan tahap *fine-tuning* (Rizky & Hidayat, 2025). Pada tahap *pre-training*, model BERT dilatih menggunakan data yang tidak memiliki label dengan berbagai tugas bahasa. Setelah itu, tahap *fine-tuning*, model BERT diinisialisasi dengan parameter yang telah diperoleh selama *pre-training*, dan semua parameter disesuaikan menggunakan data yang memiliki label (Koto dkk., 2020).

2.1.19 RoBERTa

RoBERTa merupakan singkatan dari *Robustly Optimized BERT Approach* dan dikembangkan oleh Facebook AI untuk meningkatkan kinerja BERT melalui optimasi-optimasi kunci dalam proses pelatihan. Salah satu perubahan paling signifikan dalam RoBERTa adalah penghapusan tujuan *Next Sentence Prediction* yang menyederhanakan proses pelatihan dan fokus sepenuhnya pada *masked language modelling* (M.Prytula, 2024). RoBERTa menggunakan *masking* dinamis, di mana token yang berbeda disembunyikan pada setiap *epoch*. Hal ini menghasilkan pelatihan yang lebih *robust* dan generalisasi yang lebih baik. Model ini dilatih dengan *mini-batch* yang lebih besar, tingkat

pembelajaran yang lebih tinggi, dan data yang jauh lebih banyak, mencakup 160 GB teks dari berbagai sumber (Liu dkk., 2019).

2.1.20 mBERT

Multilingual BERT (mBERT) adalah bagian dari BERT yang telah dilatih sebelumnya menggunakan data Wikipedia dalam 104 bahasa (Ragab dkk., 2025). Model ini dilatih menggunakan data Wikipedia yang seimbang berdasarkan ketersediaan sumber daya bahasa dengan menerapkan kosakata bersama untuk bahasa-bahasa yang didukung. Dalam arsitektur utamanya mBERT memiliki 12 lapisan *Transformer* dengan sekitar 86 juta parameter. Untuk kosakata yang besar, diperlukan sekitar 92 juta parameter pada lapisan *embedding* (Acs dkk., 2023).

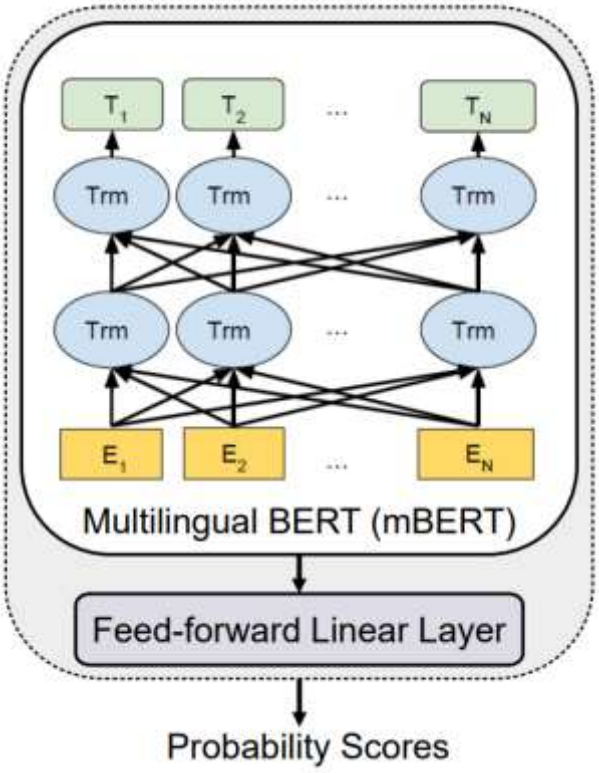
$$Attention\ (Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{2.3}$$

Keterangan:

a. Q (*Query*), merepresentasikan kata yang sedang mencari informasi.
b. K (*Key*), merepresentasikan kata lain yang akan dibandingkan tingkat keterkaitannya.
c. V (*Value*), informasi yang akan diambil jika kata tersebut dianggap penting.

Model BERT multibahasa ini mampu memproses teks dalam berbagai bahasa secara efektif (Ragab dkk., 2025). Berbagai penelitian menunjukkan bahwa *fine-tuning* mBERT pada dataset multibahasa dapat secara signifikan meningkatkan akurasi dan skalabilitas model. Selain itu, *embedding* kontekstual yang dihasilkan mBERT mampu menangkap nuansa linguistik dan budaya

(Ragab dkk., 2025). Pada tahap *fine-tuning*, mBERT dilatih ulang dengan menambahkan lapisan klasifikasi diatas arsitektur *transformer*, di mana proses tokenisasi dilakukan hingga tingkat *subword* menggunakan *WordPiece* untuk menangani kata yang tidak terdapat dalam kosakata (Ali dkk., 2023).



Gambar 2. 2 Model Multilingual BERT (mBERT) (Abdul Aziz dkk., 2022)

Model *Multilingual* BERT (mBERT) pada Gambar 2.2 menunjukkan arsitektur *transformer encoder* bertingkat yang memproses token multibahasa melalui mekanisme *self-attention bidirectional* untuk menghasilkan representasi kontekstual. Representasi ini diteruskan ke lapisan linear untuk menghasilkan skor probabilitas pada tugas klasifikasi. Arsitektur mBERT ini memungkinkan transfer pembelajaran *cross-lingual* tanpa pelatihan tambahan karena menggunakan kosakata bersama (Abdul Aziz dkk., 2022).

$$z = WT1 + b \quad (2.4)$$

Keterangan:

a. T1, representasi vektor dari token [CLS] yang dianggap sebagai rangkuman seluruh kalimat.

b. W, *Weight matrix* atau bobot model digunakan untuk memetakan fitur ke kelas emosi.

c. B, bias yaitu nilai tambahan untuk membantu penyesuaian prediksi.

d. z, *output* skor awal atau logits sebelum softmax.

Algoritma mBERT

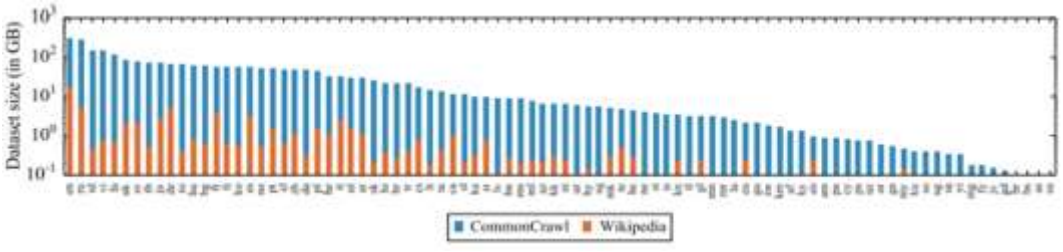
Input: Text Sequence
Start
1. Embedding Layer E1, …En
   a. Melakukan tokenisasi teks menggunakan tokenizer mBERT
      ➔ Menghasilkan input_ids dan attention_mask
   b. Mengubah setiap token menjadi vector embedding Ei
   c. Menambahkan positional embedding untuk mempertahankan urutan token
2. Transformer Layers (Trm)
   Untuk setiap Transformer layer lakukan
      Untuk setiap attention head lakukan
         a. Menghitung matriks Query (Q), Key (K), dan Value (V)
         b. Menghitung attention score menggunakan Self-attention: $Attention\ (Q,K,V) = softmax\ \left(\frac{QK^T}{\sqrt{d_k}}\right)V$
         c. Menghasilkan representasi kontekstual antar token
         Selesai
      selesai
3. Representation Output T1, …Tn
   a. Menghitung representasi kontekstual akhir
   b. Mengambil representasi token [CLS] sebagai representasi kalimat.
4. Feed-forward linear layer
   a. Memasukkan vector T1 ke lapisan klasifikasi linear: $z = WT1 + b$
   b. Memetakan representasi kelas emosi
5. Probability Calculation
   a. Menghitung probabilitas setiap kelas menggunakan softmax

$$softmax\ (z_i) = \frac{\exp\ (z_i)}{\sum_{j=1}^{v} exp\ (z_j)}$$

Algoritma mBERT
b. Menentukan kelas emosi dengan probabilitas tertinggi. **End** **Output:** Probability Scores

2.1.21 XLM-RoBERTa

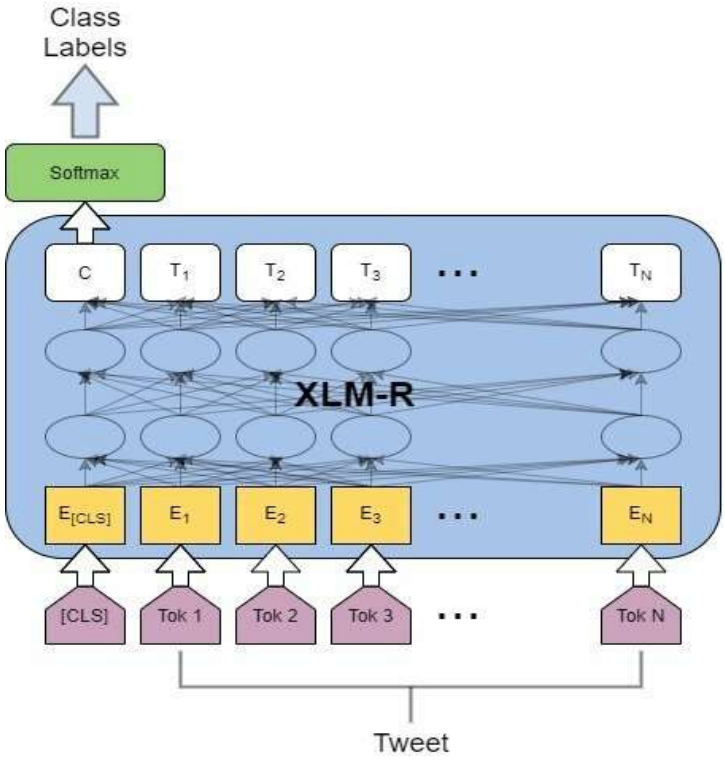
XLM-RoBERTa merupakan model multibahasa dari RoBERTa yang dikembangkan oleh Meta AI untuk memproses lebih dari 100 bahasa, menggunakan arsitektur transformer dengan mekanisme *multi-headed attention* yang mampu menangkap dependensi kata yang kompleks dan tokenisasi *Byte-Pair Encoding* (BPE) yang membuatnya lebih tahan terhadap variasi ejaan serta struktur sintaksis (Rakhimzhanov & Belginova, 2025). XLM-RoBERTa atau *Cross Lingual* Model - RoBERTa merupakan pengembangan dari XLM dan mBERT yang termasuk *multilingual Natural Language Processing* (NLP). Kemampuan XLM-RoBERTa dalam menangani ambiguitas sintaksis dan semantik terlihat jelas pada berbagai dataset campuran bahasa. Dalam bidang *cross-lingual*, XLM-RoBERTa telah menunjukkan performa yang sangat baik dengan memanfaatkan korpus paralel untuk menjembatani perbedaan antara bahasa bersumber daya tinggi dan rendah (Salve & Tubil, 2025).



Gambar 2. 3 Jumlah data pada GiB untuk 88 bahasa (Conneau dkk., 2020)

Gambar 2.3 menunjukkan pengembangan XLM-RoBERTa yang menggunakan metode *Mask Language Modeling* (MLM). MLM adalah bagian

dari *Transformer* model yang memiliki tujuan untuk memprediksi token yang hilang dari sebuah input dan merekonstruksi ulang urutan input yang sebenarnya. Namun, pada MLM tidak memiliki akses terhadap keseluruhan input tersebut, tetapi hanya memiliki akses terhadap masked token.



Gambar 2. 4 Model XLM-RoBERTa (Ramanathan dkk., 2023)

Gambar 2.4 menunjukkan model XLM-RoBERTa dapat menerima urutan masukan dengan maksimal 512 token dan menghasilkan representasi urutan. Pada proses klasifikasi dimulai dari input teks, lalu di tokenisasi menjadi beberapa token, termasuk pada token khusus [CLS] yang disimpan pada awal kalimat untuk merepresentasi setiap token. *Embedding* tersebut lalu diproses oleh lapisan encoder transformer XLM-RoBERTa yang tersusun atas mekanisme *self-attention multi-head* dan *feed-forward network*. Selanjutnya,

fungsi *softmax* yang diterapkan pada skor logit untuk mengubah menjadi distribusi probabilitas kelas sebagaimana ditunjukkan pada Persamaan (2.5).

$$softmax\ (z_i) = \frac{\exp\ (z_i)}{\sum_{j=1}^{v} exp\ (z_j)} \quad (2.5)$$

Keterangan:

a. Exp() adalah fungsi eksponensial
b. Z_i adalah nilai skor atau logit untuk token ke I dalam kosa kata yang mungkin mengisis posisi yang di [MASK]
c. Z_j adalah nilai yang menunjukkan tingkat kecocokan bahwa token ke j adalah token yang benar pada saat [MASK]
d. V adalah jumlah total token dalam kosa kata yang relevan untuk model

Fungsi *softmax* yang digunakan pada lapisan akhir model dirumuskan pada persamaan (2.5) diatas, dimana probabilitas label y_i untuk token atau input x_i dihitung berdasarkan skor logit z_i yang dihasilkan oleh model. Normalisasi melalui penjumlahan eksponensial seluruh logit memastikan untuk total probabilitas seluruh kelas bernilai 1, sehingga memungkinkan model memilih kelas yang probabilitas tertinggi sebagai hasil prediksi (Sanjaya dkk., 2025).

Algoritma XLM-RoBERTa

```
Input: Text sentence
Start
   1. Tokenisasi: [CLS], Tok1, Tok2, …, Tokn
      a. Melakukan tokenisasi menggunakan tokenizer XLM-RoBERTa
      b. Menambahkan token khusus [CLS] pada awal kalimat
      c. Menghasilkan input_ids dan attention_mask
   2. Embedding: E [CLS], E1, E2, …, En
      a. Mengubah setiap token menjadi vector embedding
      b. Menambahkan positional embedding untuk mempertahankan
         urutan token
   3. Transformer processing: XLM-RoBERTa
      untuk setiap Transformer layer lakukan
        untuk setiap attention head lakukan
```

Algoritma XLM-RoBERTa

a. Menghitung matriks Query (Q), Key (K), dan Value (V)
b. Menghitung attention score menggunakan Self-attention: $Attention\,(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$
c. Menghasilkan representasi kontekstual antar token
selesai
selesai

4. Extraction
   a. Menghasilkan representasi kontekstual akhir
   b. Mengambil vektor token [CLS] sebagai representasi kalimat
5. Klasifikasi
   a. Memasukkan vektor [CLS] ke linear layer: $z = WT1 + b$
   b. Memetakan representasi ke kelas emosi
6. Softmax
   a. Menghitung probabilitas setiap kelas emosi:

$$softmax\,(z_i) = \frac{\exp\,(z_i)}{\sum_{j=1}^{v} exp\,(z_j)}$$

7. Prediksi
   a. Memilih label emosi dengan probabilitas tertinggi

End
Output Emotions Class labels

2.1.22 Early Stopping

Early stopping adalah teknik regulasi yang digunakan dalam pelatihan model dan *deep learning* untuk mencegah *overfitting* serta hasil yang optimal (Karim & Wibowo, 2025). Teknik ini menggunakan matriks seperti *loss* atau *F1-score* selama beberapa *epoch* dan menghentikan proses pelatihan jika tidak ada lagi peningkatan (Karim & Wibowo, 2025). Pada *early stopping* juga dapat mengurangi waktu pelatihan. Teknik ini berguna karena dapat menghentikan pelatihan model segera setelah model mencapai hasil optimal pada data validasi. Selama proses pelatihan, kinerja model dievaluasi di setiap *epoch* menggunakan data validasi. *Early stopping* menggunakan parameter *patience* yang berfungsi sebagai batas toleransi jumlah *epoch* selama kinerja pada data validasi tidak meningkat. Parameter *patience* menentukan berapa kali proses pelatihan

berlanjut meskipun matriks evaluasi tidak menunjukkan peningkatan (Andika Surya dkk., 2025) .

2.1.23 *Confusion Matrix*

Confusion matrix adalah tabel evaluasi yang menggambarkan kinerja model klasifikasi dengan membandingkan prediksi model tersebut dengan label sebenarnya pada data uji. *Confusion matrix* berisi informasi mengenai jumlah klasifikasi yang benar dan salah pada data uji (Normawati & Prayogi, 2021). *confusion matrix* menunjukkan jumlah *True Positif* (TP), *True Negatif* (TN), *False Positif* (PF), dan *False Negatif* (FN) yang dihasilkan oleh model klasifikasi terlampir pada Tabel 2.1.

Tabel 2. 1 Confusion Matrix

		Kelas Prediksi	
		Positif	Negatif
Kelas sebenarnya	Positif	TP	FN
	Negatif	FP	TN

Notasi pada *Confusion* Matrix dapat dijelaskan sebagai berikut:

a. TP adalah *true positive* yaitu sebuah data dengan kondisi aktual yang mampu memprediksi dengan benar.
b. TN adalah *true negative* adalah data negatif yang diprediksi dengan benar.
c. FP adalah *false positive* yaitu sebuah data prediksi yang tidak sesuai.
d. FN adalah *false negative* yaitu sebuah data negatif yang diprediksi tidak sesuai.

Tabel 2. 2 *Multiclass Confusion Matrix*

		Kelas Prediksi			
Kelas sebenarnya		C_1	C_2	…	C_N
	C_1	$C_{1,1}$	FP	…	C_N
	C_2	FN	TP	…	FN
	…	…	…	…	…
	C_N	$C_{N,1}$	FP	…	$C_{N,N}$

Rumus yang digunakan untuk menghitung nilai akurasi model berdasarkan *confusion matrix multiclass* terdapat pada persamaan (2.6).

$$Accuracy = \frac{TF + TN}{TP + TN + FP + FN} \tag{2.6}$$

Accuracy merupakan rasio prediksi benar atau positif dan negatif dengan keseluruhan data. Matriks ini paling umum digunakan karena mudah dihitung dan digunakan, akan tetapi matriks ini memiliki kekurangan yaitu kurang akurat untuk data yang tidak seimbang (Putra dkk., 2022).

$$Precision = \frac{TP}{TP + FP} \tag{2.7}$$

Precision mengukur berapa banyak label positif yang diprediksi yang ternyata benar Persamaan (2.7).

$$Recall = \frac{TP}{TP + FN} \tag{2.8}$$

Recall juga dikenal sebagai sensitivitas mengukur berapa banyak sampel positif yang sebenarnya berhasil diprediksi dengan benar Persamaan (2.8).

$$F1 = 2\,X\,\frac{Precision\ x\ Recall}{Precision + Recall} \tag{2.9}$$

Karena *precision* dan *recall* sering kali saling bertentangan, skor F1 memberikan nilai rata-rata yang seimbang dari keduanya Persamaan (2.9).

2.2 State of The Art

Penelitian terkini mengenai klasifikasi emosi, khususnya menggunakan metode BERT. State of the art yang disajikan pada Tabel 2.3 berisi hasil penelitian yang telah dilakukan sebelumnya yang menggambarkan berbagai pendekatan yang telah diteliti untuk meningkatkan kinerja klasifikasi emosi pada teks.

Tabel 2. 3 State of The Art

No	Nama Peneliti	Nama Jurnal	Dataset	Metode	Kelebihan dan Kekurangan	Kekurangan dan Gap Penelitian
1.	(Alqarni dkk., 2025)	Emotion-Aware RoBERTa enhanced with emotion-specific attention and TF-IDF gating for fine-grained emotion recognition Fatimah	Dataset public dari Kaggle.	Emotion-Aware RoBERTa yang ditingkatkan Emotion-Specific Attention (ESA) berbasis TF-IDF.	Model Emotion-Aware RoBERTa mencapai akurasi 96.77% dan f1 0,97 pada dataset utama. Model hanya melakukan single-label classification.	Belum diuji pada bahasa campuran, arsitektur model cukup berat secara komputasi.
2.	(Hosseinzadeh dkk., 2025)	XLM-Muriel at SemEval-2025 Task 11: Hard Parameter Sharing for Multi-lingual Multi-label Emotion Detection	BRIGHTER	XLM-RoBERTa-Large dengan arsitektur hard-parameter-sharing.	Model penelitian ini mencapai F1 makro 0,84 pada *High-Resource Language* dan 0,63 pada *low-resource languages*. Encoder *multilingual* dengan *language-specific*	Perbedaan jumlah data pelatihan antar bahasa yang menyebabkan tidak konsisten.

No	Nama Peneliti	Nama Jurnal	Dataset	Metode	Kelebihan dan Kekurangan	Kekurangan dan Gap Penelitian
					expert heads untuk beberapa bahasa bisa mengalami overfit	
3.	(Fatima dkk., 2025)	NUST Titans at SemEval-2025 Task 11: AfroEmo: Multilingual Emotion Detection with Adaptive Afro-XLM-R	SemEval-2025 Task 11	adaptif pre-training dan fine-tuning model Afro-XLM-R untuk deteksi emosi multilingual.	Menunjukkan hasil pada bahasa Amharic dengan Macro-F1 sebesar 0,71, serta hasil baik pada bahasa Hausa dan Yoruba. Kesulitan dalam membedakan emosi yang mirip seperti fear dan sadness, kelas emosi yang jarang muncul seperti disgust dan surprise memiliki f1 rendah,	ketergantungan emosi tertentu pada konteks budaya, dan ketidakseimbangan kelas yang mempengaruhi performa model.
4.	(Magistra Apta Paramarta dkk., 2025)	PERFORMANCE ANALYSIS OF MULTILINGUAL AND MONOLINGUAL MODELS IN PREDICTING INDONESIAN LANGUAGE EMOTION USING	Dataset bahasa Indonesia yang terdapat dari Twitter, dan dataset bahasa Inggris yang digunakan untuk model Multilingual.	IndoBERT sebagai model monolingual, serta mBERT dan XLM-R sebagai model multilingualnya.	XLM-R mencapai F1-score sebesar 0,7793. IndoBERT dengan F1-score 0,7733 dan mBERT dengan F1-score 0,6632. Adanya kesalahan pada beberapa label emosi	Ukuran dataset bahasa Indonesia yang relatif kecil sekitar 4.000 data sehingga berdampak pada performa model.

No	Nama Peneliti	Nama Jurnal	Dataset	Metode	Kelebihan dan Kekurangan	Kekurangan dan Gap Penelitian
		TWITTER DATASET				
5.	(Aripin dkk., 2023)	Optimizing the Accuracy of the Semantic-Based Compound Emotion Classifications using the XLM RoBERTa	Dataset bahasa Indonesia yang diambil dari Twitter.	XLM-RoBERTa.	Model klasifikasi emosi berbasis XLM-RoBERTa akurasi sebesar 95,56% dan loss sebesar 12,90%. Model kurang optimal pada kalimat dengan kata unik yang memiliki makna berbeda setelah proses stemming.	Keterbatasan dataset belum mencakup variasi kata yang luas, dan model belum diuji secara luas.
6.	(Ricciardi & Manisera, 2025)	A multilingual BERT-based classification of reviews for enhanced visitor's experience analysis	Dataset dari ulasan wisata diperoleh dari Fondazione Brescia Musei.	mBERT dan XLM-R, dan clustering menggunakan algoritma HDBSCAN untuk ekstraksi kata kunci spesifik klaster.	Penelitian ini menggabungkan teknik ekstraksi kata kunci dari klaster hasil fine-tuning untuk analisis yang lebih mendalam pada ulasan wisata budaya. Tidak mempertimbangkan model generative yang lebih besar dan kompleks karena keterbatasan sumber daya komputasi,	Gap penelitian ini penggunaan model yang generative lebih besar dan kompleks.
7.	(Ragab dkk., 2025)	Multilingual Propaganda Detection: Exploring	Dataset SinaLab (2024).	mBERT, mT5, dan XLM-RoBERTa	mT5 mampu menangkap linguistik secara mendalam dengan *frame*	Gap pada penelitian ini keterbatasan representasi bahasa

No	Nama Peneliti	Nama Jurnal	Dataset	Metode	Kelebihan dan Kekurangan	Kekurangan dan Gap Penelitian
		Transformer-Based Models Mbert, XLM-RoBERTa, and Mt5			*work text-to-text* dengan akurasi 99.61%. Penerapan balancing data yang meningkatkan performa model. Ketergantungan yang berlebihan pada model individual yang menyebabkan bias dari data pretraining. Kurang eksplorasi metode ensemble.	bersumber daya data (*low-resource languages*), kebutuhan optimasi hyperparameter dan pengembangan metode ensemble yang lebih canggih, serta eksplorasi *cross-lingual transfer learning* untuk meningkatkan generalisasi dan kinerja model multibahasa.
8.	(Ali dkk., 2023)	Toward Enhanced Identification of Emotion from Resource-Constrained Language through a novel Multilingual BERT Approach	dataset Roman Urdu Dataset for Emotions (RUDE) dan Roman Urdu (RU-UCL).	Transfer learning model BERT, yaiut mBERT.	Memanfaatkan transfer learning mBERT untuk mengatasi keterbatasan data dan kompleksitas bahasa Roman Urdu, mencapai akurasi klasifikasi emosi sebesar 90%. Masih bergantung pada dataset yang kecil dan terbatas,	Gap nya perlu mengembangkan dataset lebih besar dan seimbang. Model multibahasa untuk meningkatkan generalisasi pada berbagai bahasa.

No	Nama Peneliti	Nama Jurnal	Dataset	Metode	Kelebihan dan Kekurangan	Kekurangan dan Gap Penelitian
9.	(Hosen dkk., 2025)	XLM-RoBERTa-LIME: A Novel Explainable Framework for Sentiment Analysis on Bangla-English-Hindi Code-Mixed Text	Dataset public Kaggle multilingual Bangla-Inggris-Hindi.	XLM-RoBERTa yang dikombinasikan dengan pendekatan Explainable AI menggunakan LIME (Local Interpretable Model-agnostic Explanations).	Mampu menangani teks code-mixed multibahasa, dan memberikan interpretabilitas pada prediksi model melalui LIME, mencapai akurasi 92,2%. Masih kesulitan menangani ekspresi netral atau ambigu pada teks code-mixed.	Gap pada penelitian ini, pengembangan metode yang lebih robust untuk ambiguitas emosi atau sentiment.
10.	(Hao & Liu, 2025)	Transfer Learning Approach for Sentiment Analysis in Low-Resource Austronesian Languages Using Multilingual BERT	Dataset pada media sosial bahasa Buginese dan Sundanese.	mBERT, *Cross-lingual transfer learning.*	Kombinasi data augmentation dan *cross-lingual transfer learning* pada model mBERT bahasa Austronesia, akurasi 92%. Terbatasnya generalisasi model terhadap variasi bahasa informal.	Gap pada penelitian ini perlunya pengujian pada dataset yang lebih besar, beragam, dan *cross-lingual*.

Berdasarkan analisis berbagai penelitian dapat dilihat pada Tabel 2.3 dapat disimpulkan bahwa model berbasis BERT, khususnya mBERT dan XLM-RoBERTa, memberikan hasil yang lebih baik dalam analisis emosi dan analisis sentiment teks

multilingual, termasuk teks dalam bahasa yang *low-resource.* Namun, beberapa penelitian masih memiliki keterbatasan, misalnya terkait generalisasi *cross-lingual*, kelas emosi yang tidak seimbang, kesulitan dalam menangani emosi yang ambigu, dan keterbatasan dalam skenario *cross-lingual* yang lebih luas. Oleh karena itu, penelitian lebih lanjut diperlukan untuk mengevaluasi keefektifan penyempurnaan mBERT dan XLM-RoBERTa pada kumpulan data *multilingual* dengan sejumlah label emosi, sehingga meningkatkan ketahanan dan kemampuan generalisasi model dalam klasifikasi emosi *cross-lingual*.

2.3 Matriks Penelitian

Matriks perbandingan pada berbagai penelitian terdahulu disajikan pada Tabel 2.4 merangkum berbagai studi terkait klasifikasi emosi berbasis teks, memuat informasi mengenai penulis dan tahun penelitian, dataset yang digunakan, tujuan penelitian, serta algoritma yang diterapkan, sehingga dapat memberikan gambaran terhadap perkembangan dan perbedaan pendekatan penelitian sebelumnya.

Tabel 2. 4 Matriks Penelitian

No	Penelitian	Data						Tujuan Penelitian			Algoritma					
		Indonesia	Inggris	Thai	Spanyol	Bilingual	*Multilingual*	Klasifikasi emosi	*Fine-tuning*	*Cross-lingual*	XLM-RoBERTa	BERT	IndoBERT	IndoBERTa	SVM	mBERT
1.	(Rizky & Hidayat, 2025)	√						√	√				√			
2.	(Ameer, Bölücü, Sidorov, dkk., 2023)						√	√	√		√					√
3.	(Hassan dkk., 2022)						√	√	√	√						√
4.	(RACHMAWATI, 2025)		√					√	√			√				
5.	(Kamolchanok Laojampa dkk. 2025)			√				√	√		√	√				
6.	(Magistra Apta Paramarta dkk. 2025)	√						√	√		√	√	√			
7.	(Wiciaputra dkk., 2021)					√			√	√	√					
8.	(Qu, 2021)				√			√	√		√					
9.	(Ameer dkk., 2022)		√					√				√			√	
10.	**Penelitian ini**						√	√	√	√	√					√

Berdasarkan sembilan penelitian yang tercantum dalam Tabel 2.4, dapat disimpulkan bahwa penelitian sebelumnya terutama berfokus pada klasifikasi emosi berbasis teks dengan menggunakan pendekatan *deep learning*, khususnya model *transformer* seperti BERT, mBERT, dan XLM-RoBERTa. Meskipun penelitian-penelitian ini telah memberikan kontribusi yang signifikan dalam meningkatkan akurasi dan efektivitas klasifikasi emosi, masih terdapat keterbatasan dalam menangani kompleksitas dataset *multilingual*, terutama terkait pola semantik, gaya ekspresi emosi, dan kemampuan generalisasi model

cross-lingual. Oleh karena itu, penelitian ini berfokus pada penerapan dan optimalisasi penyempurnaan model berbasis BERT, khususnya mBERT dan XLM-RoBERTa, untuk mengatasi kompleksitas teks multibahasa dan menganalisis perbedaan kinerja model sebelum dan sesudah *fine-tuning* dalam klasifikasi emosi *cross-lingual*.