

# **BAB I**

## **PENDAHULUAN**

Bab pendahuluan ini membahas mengenai latar belakang, rumusan masalah, tujuan dan manfaat penelitian, ruang lingkup penelitian, serta sistematika penulisan yang digunakan dalam skripsi yang berjudul “Analisis Sentimen Unggahan di Platform X Terkait Mangrove Menggunakan *IndoBERTweet-BiGRU* dengan Interpretasi *LIME* dan *SHAP*”.

### **1.1 Latar Belakang**

Hutan mangrove merupakan ekosistem pesisir yang memiliki peran penting dalam melindungi garis pantai dari abrasi, menjaga keanekaragaman hayati, dan penyimpanan karbon biru serta menyediakan berbagai jasa ekosistem lainnya. Namun, nilai ekologis tersebut seringkali tidak sejalan dengan persepsi masyarakat, karena mangrove sering dianggap sebagai kawasan yang kotor atau berbahaya sehingga memengaruhi upaya pelestariannya (Dahdouh-Guebas dkk., 2020). Kondisi ini semakin menantang karena tekanan aktivitas manusia seperti konversi lahan, penebangan, dan pemanfaatan sumber daya menyebabkan degradasi serta menurunnya kualitas mangrove, sehingga diperlukan pendekatan restorasi yang melibatkan partisipasi aktif masyarakat (Del Cid-Alvarado dkk., 2024). Oleh karena itu, pemahaman publik terhadap fungsi ekologis mangrove menjadi faktor penting dalam keberhasilan konservasi, karena persepsi positif dan kesadaran lingkungan terbukti meningkatkan keterlibatan dalam kegiatan perlindungan dan rehabilitasi ekosistem (Mar dkk., 2021).

Dalam konteks memahami persepsi publik, platform X hadir sebagai ruang percakapan publik yang bersifat terbuka dan *real time* sehingga efektif untuk menelusuri opini serta pola komunikasi masyarakat terhadap berbagai isu sosial maupun lingkungan. Karakteristik datanya yang organik memungkinkan analisis tren dan sentimen secara luas (Lee & Song, 2022). Selain itu, dinamika diskusi di platform X mencerminkan keberagaman respons publik karena pengguna dengan latar belakang dapat menyampaikan dukungan, kritik, maupun pandangan yang saling bertentangan terhadap suatu peristiwa, sehingga menghasilkan gambaran opini yang tidak homogen (Ali dkk., 2022). Studi lain juga menunjukkan bahwa platform X menjadi ruang penting dalam mengamati persepsi masyarakat terhadap isu lingkungan, karena berbagai kelompok seperti akademisi,

organisasi lingkungan, hingga masyarakat umum aktif membagikan informasi dan membentuk komunitas diskusi mengenai topik-topik ekologis (Bruzzese dkk., 2022). Dengan demikian, platform X menjadi sumber data yang relevan untuk menelusuri persepsi publik secara lebih luas, sekaligus menyediakan peluang untuk melakukan analisis sentimen berbasis teks terhadap isu mangrove.

Analisis sentimen terhadap unggahan publik dapat dilakukan melalui *Natural Language Processing (NLP)*, yaitu bidang kecerdasan buatan yang banyak digunakan karena memungkinkan komputer memahami, menafsirkan, dan memproses bahasa manusia melalui berbagai tahapan linguistik mulai dari fonologi sebagai analisis bunyi, morfologi sebagai pembentukan kata, sintaksis sebagai struktur kalimat, hingga semantik dan pragmatik sebagai pemaknaan dan konteks penggunaan bahasa (Khurana dkk., 2022). Dalam analisis sentimen, *NLP* berperan penting karena mampu mengekstraksi opini, emosi, serta kecenderungan sikap dari data tekstual secara otomatis, sehingga membantu memahami persepsi publik secara lebih komprehensif (Jim dkk., 2024). Akan tetapi, pendekatan *NLP* tradisional memiliki tantangan seperti ambiguitas makna, variasi gaya bahasa, serta kebutuhan untuk memahami konteks yang luas.

Perkembangan metode modern menghadirkan arsitektur *transformer* sebagai solusi melalui mekanisme *self-attention* untuk menghitung bobot relevansi antar-token sehingga model dapat menangkap hubungan semantik dan dependensi jarak jauh dalam kalimat secara lebih efektif dibandingkan pendekatan sekuensial tradisional (Alahmadi dkk., 2025). Salah satu model *transformer* yang banyak digunakan adalah *Bidirectional Encoder Representations from Transformers (BERT)* yang bekerja dengan memproses teks secara dua arah (*bidirectional*) dan merepresentasikan setiap token ke dalam vektor berdimensi tinggi untuk menangkap informasi kontekstual yang kompleks, dan berbagai penelitian menunjukkan bahwa pendekatan ini menghasilkan performa yang signifikan dalam tugas klasifikasi sentimen pada beragam bahasa, domain, dan platform media sosial karena kemampuannya menangkap nuansa emosional serta struktur bahasa yang lebih kompleks (Sayeed dkk., 2023).

Pemilihan model bahasa yang sesuai dengan karakteristik bahasa dan domain merupakan aspek krusial karena model *BERT* umum sering kali tidak optimal ketika diterapkan pada domain atau gaya bahasa yang berbeda dari data pra-pelatihannya, kinerja *BERT* terbukti meningkat secara signifikan ketika model tersebut dilatih ulang menggunakan teks yang sesuai dengan domain target (Ishigaki dkk., 2023). Pentingnya kesesuaian bahasa dan domain juga diperkuat oleh temuan bahwa model *BERT* umum mengalami *vocabulary mismatch* ketika diterapkan pada teks *Twitter* berbahasa Indonesia yang memiliki karakteristik informal seperti slang, singkatan, dan gaya bahasa khas media sosial (Koto dkk., 2021). Untuk mengatasi hal tersebut, *IndoBERTweet* dikembangkan melalui *domain-adaptive pre-training* pada lebih dari 26 juta *tweet* berbahasa Indonesia serta mengimplementasikan strategi inisialisasi kosakata yang secara teknis mampu menangani keterbatasan representasi *subword* pada teks media sosial (Koto dkk., 2021). Oleh karena itu, pada penelitian ini digunakan *IndoBERTweet* sebagai model pra-latih karena kesesuaiannya dengan bahasa Indonesia dan karakteristik domain media sosial, seperti platform X.

Pemilihan *IndoBERTweet* sebagai model pra-latih yang sesuai dengan karakteristik bahasa dan domain media sosial kemudian diikuti dengan penggunaan lapisan klasifikasi tambahan berupa *Bidirectional Gated Recurrent Unit (BiGRU)* untuk memperkuat pemrosesan urutan kata. *BiGRU* menjadi model yang relevan sebagai lapisan klasifikasi karena mampu memproses urutan kata secara dua arah melalui mekanisme *forward GRU* dan *backward GRU*, sehingga representasi yang dihasilkan dapat meningkatkan kemampuan dalam menangkap dinamika teks media sosial yang bersifat ringkas dan tidak linear dengan mempertimbangkan informasi sebelum dan sesudah suatu token (Lin & Nuha, 2023). Hal ini sangat penting untuk teks media sosial yang sering mengandung struktur kalimat tidak teratur, singkatan, dan konteks implisit yang hanya dapat dipahami jika kedua arah dependensi diperhitungkan. Selain itu, Khan dkk. (2025) juga menjelaskan bahwa setiap *unit GRU* dalam *BiGRU* memiliki *update gate* dan *reset gate* yang berfungsi mengatur aliran informasi secara adaptif, mempertahankan informasi penting, dan mengabaikan detail yang tidak relevan, yang memungkinkan *BiGRU* mempelajari pola emosional yang halus dan variatif dalam teks. Dengan demikian, penggunaan *BiGRU* dalam penelitian ini memberikan dukungan penting untuk meningkatkan kemampuan model dalam memahami struktur bahasa Indonesia pada teks media sosial.

Akan tetapi, keunggulan performa model *deep learning* seperti *IndoBERTtweet-BiGRU* menghadirkan tantangan baru karena arsitekturnya yang bersifat *black box*, yang tidak memperlihatkan secara jelas bagaimana suatu prediksi dihasilkan. Fan dkk. (2021) menyatakan bahwa jaringan saraf bekerja sebagai *black box* karena proses dan dinamika pelatihannya yang kompleks, sehingga manusia tidak dapat melihat bagaimana fitur dipilih maupun keputusan dibuat. Kentour & Lu (2021) menambahkan bahwa meskipun *deep learning* mampu mencapai performa tinggi, proses generatifnya tetap tidak dapat diuraikan secara jelas akibat struktur berlapis dan penugasan bobot yang tidak deterministik. Dalam analisis sentimen, Diwali dkk. (2024) juga menjelaskan bahwa model *deep learning* memiliki performa unggul tetapi sangat minim interpretabilitas, sehingga pengguna sulit memahami alasan prediksi suatu sentimen. Quan dkk. (2023) menambahkan bahwa rendahnya transparansi pada model *deep learning* menyebabkan hasil klasifikasi tidak mudah dipercaya. Oleh sebab itu, diperlukan pendekatan yang mampu meningkatkan transparansi model tanpa mengurangi performanya.

Oleh karena itu, pendekatan *Explainable AI* ditawarkan sebagai solusi untuk meningkatkan transparansi model kecerdasan buatan dengan menyediakan penjelasan yang dapat dipahami manusia atas prediksi yang dihasilkan (Elbasiony dkk., 2024). Dalam konteks analisis sentimen, *Explainable AI* diimplementasikan untuk menjelaskan bagaimana model menentukan apakah suatu teks bernada positif, negatif, atau netral sehingga hasilnya lebih transparan dan dapat diinterpretasikan, bukan hanya sekadar menampilkan prediksi akhir (Mabokela dkk., 2024). Dua metode yang umum digunakan dalam *Explainable AI* adalah *Local Interpretable Model-Agnostic Explanations (LIME)* dan *SHapley Additive exPlanations (SHAP)* (Perikos & Diamantopoulos, 2024). *LIME* menghasilkan penjelasan dengan cara membuat model sederhana di sekitar satu prediksi, yakni dengan memodifikasi data masukan secara lokal dan kemudian menganalisis dampaknya terhadap hasil, sehingga dapat diketahui kata-kata mana yang paling menentukan klasifikasi sentimen (Abdelwahab dkk., 2022). Sedangkan *SHAP* bekerja dengan memberikan nilai kontribusi pada setiap fitur, yang merepresentasikan sejauh mana fitur tersebut memengaruhi perbedaan antara prediksi aktual dan prediksi rata-rata, *SHAP* dapat menyoroti kata atau frasa penting yang mendorong klasifikasi sentimen (Zhang dkk., 2025).

Dengan diterapkannya analisis sentimen berbasis *NLP* menggunakan model *IndoBERTtweet-BiGRU* yang didukung pendekatan *Explainable AI* dengan metode *LIME*

dan *SHAP*, penelitian ini diharapkan tidak hanya menghasilkan prediksi sentimen yang akurat, tetapi juga menyediakan penjelasan yang dapat dipahami pengguna sehingga hasil analisis menjadi lebih informatif, dapat dipercaya, serta bermanfaat dalam mendukung pengambilan keputusan terkait isu lingkungan.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang sebelumnya, rumusan masalah dalam penelitian ini adalah bagaimana mengimplementasikan metode *Explainable AI*, yaitu *LIME* dan *SHAP* untuk memberikan interpretasi terhadap hasil klasifikasi sentimen pada model *IndoBERTweet-BiGRU* agar proses prediksi yang bersifat *black box* menjadi lebih transparan dan dapat dipahami. Dari permasalahan utama tersebut, diturunkan ke dalam permasalahan khusus yaitu:

1. Bagaimana melakukan optimasi *hyperparameter* untuk memperoleh konfigurasi terbaik bagi model *IndoBERTweet-BiGRU* dalam analisis sentimen unggahan di platform X terkait mangrove?
2. Bagaimana performa pelatihan dan generalisasi model *IndoBERTweet-BiGRU* dalam melakukan analisis sentimen unggahan di X terkait mangrove berdasarkan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*?
3. Bagaimana hasil implementasi metode *Explainable AI*, yaitu *LIME* dan *SHAP* dalam memberikan interpretasi lokal dan global terhadap hasil klasifikasi sentimen yang dihasilkan oleh model *IndoBERTweet-BiGRU*?

## 1.3 Tujuan dan Manfaat

Berdasarkan rumusan masalah yang telah ditetapkan, tujuan penelitian ini adalah sebagai berikut:

1. Menghasilkan konfigurasi *hyperparameter* dari model *IndoBERTweet-BiGRU* dalam menganalisis sentimen unggahan di platform X terkait mangrove agar model memperoleh performa terbaik dan optimal.
2. Menganalisis performa pelatihan dan mengevaluasi kemampuan generalisasi model *IndoBERTweet-BiGRU* melalui pengukuran menggunakan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*.

3. Mendapatkan transparansi terhadap hasil klasifikasi sentimen yang dihasilkan oleh model *IndoBERTweet–BiGRU* melalui penerapan metode *Explainable AI*, yaitu *LIME* dan *SHAP* baik pada tingkat lokal maupun tingkat global.

Adapun manfaat yang diharapkan dapat diperoleh dari penelitian ini adalah tersedianya model analisis sentimen unggahan di platform X terkait mangrove yang dibangun menggunakan arsitektur *IndoBERTweet–BiGRU* sehingga menghasilkan prediksi yang akurat dan sesuai dengan karakteristik media sosial berbahasa Indonesia, dan dengan integrasi menggunakan *Explainable AI* melalui *LIME* dan *SHAP* menjadikan model lebih transparan dengan menyediakan interpretasi lokal dan global, sehingga hasil klasifikasi lebih mudah dipahami, meningkatkan kepercayaan pengguna, serta mengurangi sifat *black box* pada *deep learning*. Selain itu, penelitian ini dapat menjadi referensi bagi pengembangan sistem analisis sentimen yang dapat dipertanggungjawabkan dalam mendukung pengambilan keputusan berbasis opini publik terkait mangrove.

#### 1.4 Ruang Lingkup

Ruang lingkup penelitian ini dibatasi pada empat aspek utama agar proses penelitian berjalan secara terarah dan konsisten dengan tujuan penelitian, yaitu sebagai berikut:

1. Data penelitian dikumpulkan dari platform X menggunakan kata kunci “mangrove”, “bakau”, dan “hutan bakau” pada periode 1 September 2021 hingga 30 Juni 2025 yang dilakukan dalam dua bagian, yaitu periode 1 September 2021 hingga 31 Agustus 2023 yang berasal dari penelitian sebelumnya oleh Kusumaningrum dkk. (2024), dan untuk data periode 1 September 2023 hingga 30 Juni 2025 yang dikumpulkan mandiri oleh peneliti.
2. Optimasi model dilakukan melalui pengujian berbagai konfigurasi *hyperparameter*, meliputi ukuran *batch*, jumlah *epoch*, nilai *hidden size* pada *BiGRU*, tingkat *dropout*, serta dua *learning rate* berbeda untuk *IndoBERTweet* dan *BiGRU*.
3. Evaluasi model difokuskan pada empat metrik utama, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*, untuk memantau performa pelatihan dan menilai kemampuan generalisasi model.

4. Penerapan *Explainable AI* berfokus pada metode *LIME* untuk menghasilkan interpretasi pada tingkat lokal dan *SHAP* untuk memberikan penjelasan pada tingkat global terkait hasil klasifikasi sentimen.

## 1.5 Sistematika Penulisan

Sistematika penulisan digunakan untuk memberikan gambaran dan alur yang runtut terkait penyusunan laporan skripsi yang berjudul “Analisis Sentimen Unggahan di Platform X Terkait Mangrove Menggunakan *IndoBERTweet-BiGRU* dengan Interpretasi *LIME* dan *SHAP*”. Sistematika penulisan laporan dijelaskan sebagai berikut:

### BAB I PENDAHULUAN

Bab pendahuluan menyajikan latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan terkait penerapan *Explainable AI* pada model *IndoBERTweet-BiGRU* untuk analisis sentimen terkait mangrove pada platform X.

### BAB II TINJAUAN PUSTAKA

Bab tinjauan pustaka memuat penelitian terdahulu (*state of the art*), serta teori-teori yang digunakan pada penelitian ini. Bab ini mengulas beberapa terminologi antara lain analisis sentimen, *data crawling*, *few-shot learning*, teknik pra-pemrosesan, arsitektur *Transformer*, *BERT*, *IndoBERTweet*, *BiGRU*, serta *Bayesian Optimization*, metrik evaluasi, dan metode *Explainable AI*.

### BAB III METODOLOGI PENELITIAN

Bab metodologi penelitian menyajikan alur dan langkah-langkah yang digunakan sebagai acuan dalam pelaksanaan penelitian. Gambaran umum dari penelitian ini adalah pengumpulan data, pelabelan data, pra-pemrosesan data, pembagian dataset, penanganan *imbalanced dataset*, tokenisasi, pelatihan model *IndoBERTweet-BiGRU*, evaluasi, penerapan *Explainable AI* dengan *LIME* dan *SHAP*, dan *experimental setup*.

## BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi pembahasan dari seluruh langkah penelitian dan penjabaran hasil dan analisis eksperimen yang didapatkan dari pelatihan model analisis sentimen menggunakan arsitektur *IndoBERTweet–BiGRU* serta penerapan metode *Explainable AI* sesuai dengan metodologi penelitian yang dijadikan acuan.

## BAB V PENUTUP

Bab penutup berisi kesimpulan terhadap hasil dan analisis dari eksperimen yang telah dilakukan pada penelitian ini terkait pengembangan model analisis sentimen serta penerapan *Explainable AI* dan saran yang dibutuhkan untuk penelitian selanjutnya.