

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sepak bola merupakan olahraga paling populer di Indonesia dan telah berkembang menjadi fenomena sosial yang melampaui sekadar pertandingan di lapangan. Dukungan masyarakat terhadap Tim Nasional Sepak Bola Indonesia (Timnas Indonesia) tidak hanya berbentuk antusiasme pada saat pertandingan, tetapi juga muncul dalam bentuk partisipasi aktif dalam diskusi publik, kritik, dan opini mengenai dinamika internal tim, termasuk manajemen pelatih. Tingginya keterlibatan emosional masyarakat terhadap performa Timnas membuat keputusan strategis seperti pergantian pelatih menjadi isu sensitif yang memantik respons luas dari publik (Artamevia & Wibowo, 2025; Riska et al., 2025) .

Berdasarkan hasil survei oleh Ipsos (2022), Indonesia tercatat sebagai negara dengan proporsi penggemar sepak bola terbesar di dunia. Survei tersebut menunjukkan bahwa sekitar 69% masyarakat Indonesia menyukai sepak bola, angka yang lebih tinggi dibandingkan negara lain seperti Arab Saudi (67%), Uni Emirat Arab (65%), serta India (60%). Secara global, rata-rata proporsi penggemar sepak bola hanya mencapai sekitar 39%, sehingga menunjukkan bahwa minat masyarakat Indonesia terhadap olahraga ini berada jauh di atas rata-rata dunia. Survei ini dilakukan pada periode 26 Agustus hingga 9 September 2022 di 34 negara dengan total 22.528 responden berusia 18–74 tahun.

Perubahan pelatih dalam sepak bola sering dianggap sebagai langkah strategis federasi atau organisasi olahraga untuk memperbaiki performa tim, meningkatkan kualitas permainan, serta menyesuaikan gaya taktik dengan kebutuhan kompetisi (Artamevia & Wibowo, 2025; Sousa et al., 2024). Namun, perubahan tersebut tidak selalu diterima secara positif oleh masyarakat. Faktor seperti prestasi pelatih sebelumnya, rekam jejak, gaya bermain, hingga keputusan teknis yang pernah dilakukan dapat memengaruhi penilaian publik. Di Indonesia, perdebatan publik mengenai performa pelatih Timnas sangat umum terjadi, terutama ketika hasil pertandingan dianggap tidak sesuai ekspektasi masyarakat. Hal ini menunjukkan bahwa persepsi publik memiliki pengaruh sosial yang

signifikan terhadap citra pelatih maupun federasi yang membuat keputusan (Nugraha & Zakiyah, 2025; Riska et al., 2025).

Perkembangan teknologi informasi, khususnya media sosial, semakin memperbesar peran masyarakat dalam menyuarakan pendapat. Media sosial, sebagai salah satu platform komunikasi cepat yang menampung berbagai opini real-time, menjadi ruang utama bagi masyarakat untuk mengekspresikan sentimen mereka terhadap isu olahraga nasional (Drus & Khalid, 2019; Nugraha & Zakiyah, 2025; Rodríguez-Ibáñez et al., 2023). Pergantian pelatih Timnas Indonesia pada periode 2025–2026 menjadi salah satu isu yang memicu tingginya aktivitas publik di media sosial. Patrick Kluivert resmi diberhentikan oleh PSSI pada 16 Oktober 2025 melalui mekanisme mutual termination setelah gagal membawa Timnas Indonesia lolos ke Piala Dunia 2026. Selanjutnya, John Herdman resmi diumumkan sebagai pelatih baru Timnas Indonesia pada 3 Januari 2026. Perubahan tersebut memunculkan berbagai respons masyarakat dalam bentuk unggahan, komentar, tagar, dan percakapan publik yang meningkat secara signifikan di platform seperti Instagram, YouTube, Facebook, dan X/Twitter. Informasi yang terkandung dalam unggahan tersebut dapat mencerminkan dukungan, kritik, dan reaksi emosional secara langsung dari masyarakat (Birjali et al., 2021; Zafarani et al., 2014). Tingginya aktivitas publik tersebut menunjukkan bahwa isu pergantian pelatih Timnas Indonesia berkembang menjadi fenomena sosial yang relevan untuk dianalisis menggunakan pendekatan analisis sentimen berbasis *Natural Language Processing* (NLP) (Drus & Khalid, 2019; Prima Ertansyah et al., 2025). Oleh karena itu, sentimen publik menjadi aspek penting yang dapat digunakan untuk mengevaluasi dampak sosial dari keputusan perubahan pelatih Timnas Indonesia.

Untuk memahami fenomena tersebut secara ilmiah, analisis sentimen berbasis *Natural Language Processing* (NLP) menjadi metode yang tepat. NLP memungkinkan peneliti menganalisis opini publik dalam jumlah besar secara otomatis dan sistematis (Birjali et al., 2021; Prima Ertansyah et al., 2025). Salah satu model NLP paling berkembang saat ini adalah *BERT (Bidirectional Encoder Representations from Transformers)*, yang dikembangkan oleh Google dan dikenal memiliki akurasi tinggi dalam memahami konteks bahasa secara dua arah

(Devlin et al., 2019; Mudding, 2024). *BERT* terbukti lebih baik dalam menangkap makna kalimat dibandingkan metode tradisional seperti *Support Vector Machine* (González-Carvajal & Garrido-Merchán, 2020; Sabrina et al., 2025).

Di Indonesia, model *BERT* telah diadaptasi menjadi *IndoBERT* dan variasinya yang disesuaikan dengan karakter linguistik bahasa Indonesia sehingga mampu menghasilkan analisis sentimen yang lebih akurat (Koto et al., 2020; Wilie et al., 2020). Dengan memanfaatkan model *BERT* Bahasa Indonesia, analisis sentimen publik terhadap perubahan pelatih Timnas Indonesia dapat dilakukan secara komprehensif. Model ini memungkinkan klasifikasi opini menjadi sentimen positif, negatif, maupun netral dengan mempertimbangkan konteks kalimat, gaya bahasa, serta struktur linguistik khas pengguna media sosial di Indonesia (Jayadianti et al., 2022; Wilie et al., 2020).

Selain mengetahui kecenderungan sentimen positif, negatif, dan netral, analisis terhadap komentar negatif juga penting dilakukan untuk memahami arah kritik publik secara lebih spesifik. Komentar negatif dalam konteks perubahan pelatih Timnas Indonesia tidak selalu ditujukan secara langsung kepada pelatih baru, tetapi dapat pula mengarah kepada federasi, pelatih sebelumnya, keputusan pergantian pelatih, maupun pihak lain yang dianggap berpengaruh terhadap dinamika Timnas Indonesia. Oleh karena itu, penelitian ini tidak hanya melakukan klasifikasi sentimen, tetapi juga melakukan identifikasi target sentimen negatif menggunakan pendekatan *Aspect-Based Sentiment Analysis* (ABSA) secara terbatas (Pontiki & Pavlopoulos, 2014; Zhang et al., 2022). Pendekatan ini digunakan untuk mengelompokkan komentar negatif berdasarkan target kritik utama, yaitu Federasi, Pelatih Baru, Pelatih Lama, dan Lainnya.

Hasil analisis tersebut diharapkan dapat memberikan gambaran objektif mengenai persepsi masyarakat terhadap perubahan pelatih Timnas Indonesia, baik dari sisi kecenderungan sentimen maupun arah kritik negatif publik. Dengan demikian, penelitian ini tidak hanya menunjukkan apakah respons masyarakat cenderung positif, negatif, atau netral, tetapi juga memberikan informasi mengenai pihak atau aspek yang paling banyak menjadi sasaran kritik dalam isu perubahan pelatih Timnas Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian mengenai analisis sentimen publik terhadap perubahan pelatih Timnas Indonesia menggunakan model *BERT* Bahasa Indonesia, maka rumusan masalah dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimana proses pengumpulan, *cleaning*, *preprocessing*, dan pelabelan data komentar media sosial terkait perubahan pelatih Timnas Indonesia?
2. Bagaimana penerapan model IndoBERT dalam melakukan klasifikasi sentimen publik terhadap perubahan pelatih Timnas Indonesia?
3. Bagaimana performa model IndoBERT dibandingkan dengan metode pembandingan, serta bagaimana distribusi target sentimen negatif berdasarkan pendekatan ABSA terbatas?

1.3 Batasan Masalah

Agar penelitian ini terarah dan tidak melebar pada aspek yang berada di luar fokus analisis, maka penelitian ini memiliki batasan sebagai berikut:

1. Sumber dan Jenis Data

Data yang digunakan dalam penelitian ini berupa komentar media sosial berbahasa Indonesia yang berkaitan dengan pengumuman dan pengenalan John Herdman sebagai pelatih baru Timnas Indonesia. Data diperoleh dari Instagram, YouTube, X/Twitter, dan Facebook, dengan Instagram sebagai sumber utama karena memiliki jumlah interaksi yang paling tinggi. Penelitian ini hanya menganalisis teks komentar, tidak mencakup gambar, video, jumlah like, jumlah share, profil pengguna, maupun metadata lain di luar teks komentar.

2. Kategori Sentimen

Klasifikasi sentimen dalam penelitian ini dibatasi pada tiga kategori, yaitu positif, negatif, dan netral. Penelitian ini tidak melakukan klasifikasi emosi secara lebih rinci seperti marah, sedih, kecewa, senang, atau takut.

3. Pelabelan Data

Pelabelan awal dilakukan menggunakan pendekatan *Lexicon-Based Sentiment Analysis* sebagai label awal atau weak label. Label tersebut tidak digunakan sebagai label final, melainkan digunakan untuk membantu proses pengambilan sampel. Data sampel sebanyak 3.000 komentar kemudian diverifikasi dan dikoreksi secara manual berdasarkan pedoman pelabelan sentimen.

4. Model Klasifikasi Sentimen

Model utama yang digunakan dalam penelitian ini adalah IndoBERT. Metode *Support Vector Machine* (SVM) dan *Naive Bayes* digunakan sebagai model pembanding, sedangkan IndoBERT dengan *Class Weight* digunakan sebagai eksperimen tambahan. Model final yang digunakan adalah IndoBERT tanpa *Class Weight* berdasarkan hasil evaluasi dan pertimbangan kemampuan model dalam memahami konteks komentar.

5. Analisis Target Sentimen Negatif

Analisis lanjutan hanya dilakukan pada komentar yang berlabel negatif dengan pendekatan *Aspect-Based Sentiment Analysis* (ABSA) secara terbatas. ABSA dalam penelitian ini tidak digunakan sebagai ABSA penuh pada seluruh komentar, tetapi hanya untuk mengidentifikasi target sentimen negatif yang dibagi menjadi empat kategori, yaitu Federasi, Pelatih Baru, Pelatih Lama, dan Lainnya.

1.4 Tujuan Penelitian

Penelitian ini dilakukan untuk menganalisis sentimen publik terhadap perubahan pelatih Timnas Indonesia melalui pendekatan *Natural Language Processing* (NLP) dengan model IndoBERT. Secara khusus, tujuan penelitian ini adalah sebagai berikut:

1. Mengumpulkan, membersihkan, melakukan *preprocessing*, serta melakukan pelabelan data komentar media sosial terkait perubahan pelatih Timnas Indonesia.

2. Menerapkan model IndoBERT dalam melakukan klasifikasi sentimen publik terhadap perubahan pelatih Timnas Indonesia ke dalam kategori positif, negatif, dan netral.
3. Mengevaluasi performa model IndoBERT dengan metode pembandingan serta mengidentifikasi distribusi target sentimen negatif menggunakan pendekatan *ABSA* terbatas.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis maupun praktis dalam bidang analisis sentimen, pengolahan bahasa alami, serta kajian opini publik pada isu olahraga nasional. Adapun manfaat penelitian ini adalah sebagai berikut:

1. Manfaat bagi pengembangan ilmu pengetahuan

Penelitian ini dapat memperkaya kajian di bidang *Natural Language Processing (NLP)*, khususnya penerapan model *IndoBERT* dalam klasifikasi sentimen komentar media sosial berbahasa Indonesia. Selain itu, penelitian ini juga memberikan contoh penerapan analisis target sentimen negatif berbasis *ABSA* terbatas pada komentar publik yang bersifat informal dan kontekstual.

2. Manfaat bagi penelitian selanjutnya

Penelitian ini dapat menjadi referensi bagi penelitian berikutnya yang membahas analisis sentimen media sosial, penggunaan *Lexicon-Based Sentiment Analysis* sebagai label awal, verifikasi manual label, serta perbandingan model *Machine Learning* dan *Deep Learning* seperti *SVM*, *Naive Bayes*, dan *IndoBERT*.

3. Manfaat bagi federasi atau pengelola sepak bola

Hasil penelitian ini dapat memberikan gambaran berbasis data mengenai kecenderungan sentimen publik terhadap perubahan pelatih Timnas Indonesia. Selain itu, hasil identifikasi target sentimen negatif dapat membantu menunjukkan pihak atau aspek yang paling banyak menjadi

sasaran kritik publik, seperti federasi, pelatih baru, pelatih lama, atau aspek lainnya.

4. Manfaat bagi masyarakat dan pengamat sepak bola

Penelitian ini dapat memberikan informasi mengenai pola respons publik terhadap isu perubahan pelatih Timnas Indonesia secara lebih objektif. Dengan demikian, opini yang berkembang di media sosial dapat dilihat tidak hanya sebagai komentar individual, tetapi juga sebagai data yang dapat dianalisis untuk memahami kecenderungan persepsi masyarakat.

5. Manfaat bagi peneliti

Penelitian ini memberikan pengalaman dalam menerapkan tahapan analisis sentimen secara sistematis, mulai dari pengumpulan data, *preprocessing* teks, pelabelan awal berbasis *lexicon*, verifikasi manual, pelatihan model klasifikasi, evaluasi model, hingga analisis target sentimen negatif.

1.6 Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah metode kuantitatif dengan pendekatan text mining dan *Machine Learning*. Penelitian dilakukan dengan mengolah data komentar media sosial berbahasa Indonesia yang berkaitan dengan perubahan pelatih Timnas Indonesia. Data tersebut dianalisis menggunakan model IndoBERT untuk mengklasifikasikan sentimen ke dalam kategori positif, negatif, dan netral. Selain itu, penelitian ini juga melakukan analisis lanjutan terhadap komentar berlabel negatif untuk mengidentifikasi target sentimen negatif menggunakan pendekatan *Aspect-Based Sentiment Analysis* (ABSA) secara terbatas. Adapun tahapan metode penelitian yang dilakukan adalah sebagai berikut:

1. Pengumpulan Data

Tahap ini dilakukan dengan mengumpulkan komentar dari media sosial, yaitu Instagram, YouTube, X/Twitter, dan Facebook, yang berkaitan dengan pengumuman dan pengenalan John Herdman sebagai pelatih baru Timnas Indonesia. Pengambilan data dilaksanakan memanfaatkan

kombinasi alat ekstraksi Chrome IG Exporter & Scraper khusus untuk platform Instagram serta layanan cloud scraper berbasis Apify untuk platform YouTube, X/Twitter, dan Facebook. Data yang diperoleh masih berupa data mentah sehingga perlu melalui tahap pembersihan dan pengolahan lebih lanjut.

2. *Preprocessing* dan Pelabelan Awal Data

Tahap ini meliputi proses *cleaning*, case folding, normalisasi bahasa tidak baku atau slang, penghapusan duplikasi, serta filtering komentar yang terlalu pendek atau tidak relevan. Setelah data dibersihkan, dilakukan pelabelan awal menggunakan pendekatan *Lexicon-Based Sentiment Analysis*. Label yang dihasilkan pada tahap ini hanya digunakan sebagai label awal atau weak label, bukan sebagai label final.

3. Sampling dan Verifikasi Manual Label Sentimen

Dari data yang telah diberi label awal, diambil sampel sebanyak 3.000 komentar secara seimbang berdasarkan label awal *lexicon*. Sampel tersebut kemudian diverifikasi dan dikoreksi secara manual berdasarkan pedoman pelabelan sentimen. Tahap verifikasi manual dilakukan untuk meningkatkan akurasi label sehingga data yang digunakan pada proses pelatihan dan evaluasi model lebih representatif terhadap sentimen sebenarnya. Data hasil verifikasi manual digunakan sebagai data utama untuk pelatihan dan evaluasi model.

4. Pelatihan dan Evaluasi Model Klasifikasi Sentimen

Data manual sebanyak 3.000 komentar dibagi menjadi data *training* (64%), *validation* (16%), dan *testing* (20%). Pembagian ini menggunakan skema hold-out 80:20 yang diturunkan kembali secara proporsional untuk memisahkan data latih dan validasi, di mana data validasi berfungsi memantau loss dan mencegah overfitting saat *Fine-Tuning* IndoBERT. Selanjutnya, dilakukan pelatihan dan evaluasi terhadap model pembanding (SVM dan *Naive Bayes*) serta model utama (IndoBERT dan IndoBERT dengan *Class Weight*) menggunakan metrik *Accuracy*, *Precision*, *Recall*, *F1-Score*, dan *Confusion Matrix*.

5. Pemilihan Model Final dan Prediksi Data Sisa

Model final dipilih berdasarkan hasil evaluasi serta pertimbangan kemampuan model dalam memahami konteks komentar media sosial. Model IndoBERT tanpa *Class Weight* kemudian dilatih ulang menggunakan seluruh data manual sebanyak 3.000 komentar dan digunakan untuk memprediksi sentimen pada data komentar yang belum dilabeli manual. Hasil prediksi tersebut digabungkan dengan data manual untuk memperoleh dataset akhir.

6. Analisis Sentimen dan Target Sentimen Negatif

Tahap akhir dilakukan dengan menganalisis distribusi sentimen positif, negatif, dan netral pada dataset akhir. Komentar yang berlabel negatif kemudian dianalisis lebih lanjut menggunakan pendekatan ABSA terbatas untuk mengidentifikasi target sentimen negatif. Target tersebut dibagi menjadi empat kategori, yaitu Federasi, Pelatih Baru, Pelatih Lama, dan Lainnya.

1.7 Sistematika Penulisan

Sistematika penulisan disusun untuk memberikan gambaran mengenai susunan dan isi pembahasan pada tugas akhir ini sehingga memudahkan pembaca dalam memahami alur penelitian. Adapun sistematika penulisan tugas akhir ini adalah sebagai berikut.

BAB 1 : PENDAHULUAN

Bab ini berisi latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, serta sistematika penulisan. Bab ini menjadi dasar yang menjelaskan alasan, ruang lingkup, dan arah penelitian mengenai analisis sentimen publik terhadap perubahan pelatih Timnas Indonesia menggunakan model IndoBERT.

BAB II : TINJAUAN PUSTAKA

Bab ini memuat teori-teori yang mendukung penelitian, meliputi penelitian terdahulu, analisis sentimen, *Lexicon-Based Sentiment Analysis*, *Aspect-Based Sentiment Analysis (ABSA)*, *Natural Language Processing (NLP)*, klasifikasi

teks, media sosial sebagai sumber data, *BERT* dan *IndoBERT*, TF-IDF, *Naive Bayes*, *Support Vector Machine (SVM)*, *preprocessing* teks dan tokenisasi, *Class Weight*, evaluasi model klasifikasi, serta flowchart.

BAB III : METODOLOGI PENELITIAN

Bab ini menjelaskan tahapan penelitian yang dilakukan, mulai dari identifikasi masalah dan studi literatur, pengumpulan data, *cleaning* dan *preprocessing* teks, pelabelan awal menggunakan *Lexicon-Based Sentiment Analysis*, sampling dan verifikasi manual label sentimen, tokenisasi dan pembagian data, pelatihan model klasifikasi sentimen, pengujian dan evaluasi model, pemilihan model final, *training* ulang model final dan prediksi data sisa, identifikasi target sentimen negatif berbasis ABSA, hingga analisis distribusi sentimen dan target sentimen negatif.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi setiap tahapan penelitian beserta pembahasannya, meliputi hasil pengumpulan data, *cleaning* dan *preprocessing* data, pelabelan awal, verifikasi manual, tokenisasi dan pembagian data, pelatihan model, pengujian dan evaluasi model, pemilihan model final, prediksi data sisa, analisis distribusi sentimen akhir, identifikasi target sentimen negatif, serta pembahasan hasil penelitian.

BAB V : PENUTUP

Bab ini berisi kesimpulan yang diperoleh berdasarkan hasil penelitian serta saran yang dapat digunakan sebagai masukan untuk pengembangan penelitian selanjutnya.