

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sepak bola merupakan olahraga yang paling banyak penggemarnya di dunia, termasuk di Indonesia. Minat masyarakat terhadap sepak bola sangat besar, yang terbukti dari besarnya jumlah penonton yang menyaksikan pertandingan, baik secara langsung maupun melalui tayangan televisi (Hakim, 2020). Banyak negara terus meningkatkan kualitas tim nasionalnya guna memenuhi ekspektasi penggemar dan meningkatkan prestasi. Tekanan untuk tampil kompetitif diajang internasional pun turut mendorong Persatuan Sepak Bola Seluruh Indonesia (PSSI) menerapkan program naturalisasi. Program ini dirancang untuk memperkuat tim dengan menghadirkan para pemain berdarah Indonesia yang bermain di luar negeri (Sania dkk., 2025). Naturalisasi merupakan proses perolehan kewarganegaraan oleh individu asing melalui permohonan resmi kepada pemerintah negara yang ingin ditinggalinya. Namun, permohonan itu hanya bisa dikabulkan jika pemohon telah memenuhi seluruh syarat yang tercantum dalam Undang-Undang Nomor 12 Tahun 2006 mengenai Kewarganegaraan Republik Indonesia (Annas & Hazzar, 2023). Sejak tahun 2023 hingga 2025, program ini telah menghasilkan 20 pemain naturalisasi. Penambahan pemain-pemain naturalisasi semakin memperkuat Timnas Indonesia, tetapi juga menimbulkan kekhawatiran terhadap regenerasi pemain lokal dan perkembangan sepak bola nasional.

Meskipun proses naturalisasi telah dilakukan di Indonesia sejak lama, meningkatnya jumlah pemain naturalisasi dalam tim nasional saat ini menjadi fenomena baru yang memicu perdebatan di kalangan masyarakat (Zahran dkk., 2024). Perbincangan di media sosial menunjukkan adanya pro dan kontra terkait kebijakan ini. Seperti yang terjadi saat ini, program naturalisasi yang dilaksanakan oleh PSSI menerima sejumlah kecaman dari sebagian penggemar sepak bola di Indonesia. Identitas nasional menjadi isu yang pertama kali dibahas dalam kritik tersebut. Ada pendapat yang mengatakan bahwa kehadiran pemain naturalisasi bisa memudahkan identitas tim nasional yang seharusnya dihuni oleh para atlet lokal (Alam & Gono, 2024). Kehadiran pemain asing dalam tim nasional juga dipandang sebagai indikasi kurangnya dukungan negara terhadap potensi pemain lokal.

Sebagian pihak mengkritik PSSI karena dianggap terlalu agresif dalam melakukan naturalisasi, seperti yang diungkapkan dalam laporan Viva Bola, di mana seorang pengamat sepak bola menilai bahwa kebijakan ini mengesampingkan peran Liga 1 sebagai wadah pembinaan pemain lokal untuk tim nasional (Aditya & Karami, 2024). Namun, di sisi lain, beberapa orang juga berpendapat bahwa program ini telah sukses dalam meningkatkan kinerja tim nasional Indonesia. Hal ini terbukti dari pencapaian tim nasional yang berhasil lolos ke ronde ketiga babak Kualifikasi Piala Dunia 2026 dan lolos ke Piala Asia dalam dua edisi berturut-turut (Prayugi, 2024). Hal ini menunjukkan bahwa meskipun terdapat perdebatan, naturalisasi telah membawa dampak positif terhadap pencapaian Timnas Indonesia dalam beberapa turnamen internasional.

Dalam waktu beberapa tahun terakhir, kemajuan dalam bidang informasi dan komunikasi telah mendorong cepatnya pertumbuhan platform sosial, dengan X sebagai salah satu yang terfavorit di Indonesia. Sekitar 59% masyarakat Indonesia menggunakan X, menjadikannya platform berpengaruh dalam diskusi publik. Popularitasnya didorong oleh kemudahan pengguna dalam menyampaikan opini terkait berbagai topik (Utomo dkk., 2023). Dengan jutaan pengguna aktif, X berfungsi sebagai saluran bagi individu untuk menyampaikan pendapat, membagikan informasi, dan terlibat dalam diskusi tentang berbagai isu, termasuk olahraga. Salah satu hal yang menarik perhatian adalah cara pemain menjadi warga negara dalam tim nasional sepak bola Indonesia. Platform seperti X telah menjadi tempat utama bagi pecinta sepak bola untuk menyampaikan pandangan mereka tentang naturalisasi. Beragam komentar, baik yang mendukung maupun yang menolak, tentang pemain naturalisasi ditimnas Indonesia dapat ditemukan di berbagai akun media sosial X (Ramdani & Bustomi, 2024). Untuk memahami bagaimana opini publik menilai kebijakan ini, pendekatan analisis sentimen diterapkan. Menganalisis sentimen publik mengenai naturalisasi timnas Indonesia perlu dilakukan untuk memahami seberapa besar penerimaan masyarakat terhadap kebijakan tersebut (Lubis dkk., 2024).

Metode analisis sentimen adalah pendekatan dalam ilmu data yang digunakan untuk memahami dan mengelompokkan pendapat masyarakat di berbagai platform digital, seperti media sosial, forum, dan portal berita. Analisis sentimen merupakan

proses untuk menafsirkan makna dari pendapat, pandangan, atau ekspresi emosional dalam sebuah teks, percakapan, postingan online, atau basis data lainnya menggunakan *Natural Language Processing* (NLP). Proses ini bertujuan untuk mengenali dan memproses data berupa teks secara otomatis dengan tujuan memperoleh informasi yang signifikan serta mengidentifikasi opini mengenai suatu tema atau entitas tertentu didalam sekumpulan data (Hasri & Alita, 2022). Metode utamanya melibatkan konstruksi kamus data yang mengklasifikasikan kata-kata menjadi kategori positif dan negatif berdasarkan makna intrinsiknya. Selain itu, berbagai algoritma machine learning sering dipakai untuk meningkatkan ketepatan analisis sentimen, seperti *Naïve Bayes*, *Support Vector Machine* (SVM), dan *Logistic Regression*. Algoritma *deep learning*, misalnya *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM), juga digunakan untuk menangkap konteks dan pola dalam teks secara lebih mendalam.

Metode-metode yang telah dijelaskan sebelumnya sering dipakai dalam analisis sentimen pada studi-studi sebelumnya. Salah satu penelitian melakukan analisis sentimen terhadap unggahan pengguna pada media sosial X yang berkaitan dengan Pemilihan Umum Presiden Indonesia 2024 dengan memanfaatkan metode *Convolutional Neural Network* (CNN). Dataset yang digunakan berjumlah 10.250 *tweet* yang diperoleh melalui Twitter API dalam rentang waktu Juni hingga September 2024. Penelitian tersebut menguji enam variasi konfigurasi model, di mana performa terbaik dicapai oleh kombinasi CNN dengan pembobotan TF-IDF pada *epoch* ke-15. Model ini mencatat akurasi 84%, dengan nilai precision, recall, dan F1-score masing-masing berada pada angka 83% (Nurfauziyah dkk., 2024). Selain itu, terdapat penelitian lain yang mengkaji penerapan metode *Convolutional Neural Network* (CNN) dalam menganalisis sentimen terhadap opini masyarakat mengenai isu lingkungan hidup di Kota Depok melalui Twitter juga telah dilaksanakan. Penelitian ini menghasilkan akurasi sebesar 86%. Akurasi tersebut menunjukkan bahwa metode CNN mampu melakukan analisis sentimen dengan baik (Parameswari & Prihandoko, 2022).

Analisis sentimen melalui metode klasifikasi membutuhkan strategi untuk mengekstrak fitur demi meningkatkan kinerja model klasifikasi tersebut. Seperti penelitian yang dilakukan oleh Jonathan dkk, dalam penelitian tersebut dilakukan

klasifikasi teks berita berbahasa Indonesia dengan memanfaatkan IndoBERT *Embedding* serta CNN. Penelitian tersebut menggunakan dataset berupa 3.000 artikel berita yang dikumpulkan dari portal berita Kompas.com. Berdasarkan hasil pengujian dengan memanfaatkan IndoBERT sebagai representasi kata dan CNN untuk proses klasifikasi, diperoleh tingkat akurasi mencapai 91% (Tantoro & Darmawan, 2025). Penelitian ini menunjukkan bahwa kombinasi IndoBERT sebagai representasi teks dan CNN untuk klasifikasi efektif dalam analisis sentimen.

Berdasarkan berbagai penelitian sebelumnya, metode *Convolutional Neural Network* (CNN) terbukti memiliki kinerja yang baik dalam analisis sentimen, terutama pada dataset yang berasal dari media sosial seperti X. Meski CNN awalnya dirancang untuk analisis gambar, perkembangan waktu telah membuat CNN menjadi model yang efisien dan cepat dalam menangani berbagai tugas dalam pemrosesan Bahasa alami. CNN dapat menangkap pola-pola penting dalam urutan kata, seperti frasa atau kombinasi kata yang memiliki makna emosional tertentu, melalui penerapan tiga jenis lapisan yang terdapat dalam CNN, yaitu *convolutional layer*, *pooling layer*, dan *fully connected layer*. Dibandingkan dengan model tradisional seperti *Naïve Bayes* atau SVM, CNN mampu mengenali hubungan kata dalam konteks yang lebih luas tanpa harus bergantung pada fitur yang ditentukan secara manual. Selain itu, IndoBERT *Embedding* dapat meningkatkan efektivitas metode klasifikasi, yang terbukti melalui hasil test dan tingkat ketepatan model dalam penelitian (Amalia & Winarko, 2021). IndoBERT adalah versi BERT yang disesuaikan untuk Bahasa Indonesia dan dilatih menggunakan kumpulan teks besar berbahasa Indonesia. Dengan menggunakan IndoBERT *embedding*, representasi kata dalam teks dapat lebih sesuai dengan karakteristik unik bahasa Indonesia, seperti kata-kata majemuk, afiksasi, atau perbedaan makna berdasarkan konteks (Merdiansah dkk., 2024). Oleh karena itu, penelitian ini akan menerapkan metode IndoBERT *Embedding* dan CNN untuk menghasilkan representasi vektor kata yang mempertimbangkan konteks serta makna dalam teks tweet di media.

Dengan mempertimbangkan urgensi isu pemain naturalisasi dalam tim nasional sepak bola Indonesia dan efektivitas metode analisis yang telah dijelaskan, dilakukan sebuah penelitian berjudul “Analisis Sentimen Program Naturalisasi Tim Nasional Sepak Bola Indonesia Pada Media Sosial X Menggunakan IndoBERT

Embedding Dan Convolutional Neural Network”. Penelitian ini bertujuan untuk memberikan gambaran mengenai opini publik terkait kebijakan naturalisasi melalui analisis sentimen, serta menguji efektivitas kombinasi metode CNN dan IndoBERT *Embedding* dalam mengolah data teks dari media sosial. Diharapkan, temuan dari penelitian ini bisa menjadi referensi yang berguna untuk memahami respons publik terhadap kebijakan naturalisasi pemain dalam dunia sepak bola Indonesia.

1.2 Identifikasi Masalah Penelitian

Mengacu pada latar belakang yang telah diuraikan, dapat dirumuskan sejumlah permasalahan yang menjadi landasan dalam pelaksanaan penelitian ini, yaitu:

1. Meningkatnya jumlah pemain naturalisasi dalam Tim Nasional Sepak Bola Indonesia menimbulkan perdebatan publik, khususnya terkait dampaknya terhadap peluang pemain lokal, identitas nasional, dan pembinaan jangka panjang sepak bola Indonesia.
2. Belum banyak kajian ilmiah yang secara sistematis menganalisis opini masyarakat terhadap kebijakan naturalisasi.

1.3 Rumusan Masalah

Berdasarkan penjelasan mengenai masalah yang mendasari penelitian ini yang telah dipaparkan, maka perumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana kecenderungan sentimen pengguna media sosial X terhadap program naturalisasi tim nasional sepak bola Indonesia berdasarkan hasil distribusi kelas dari proses pelabelan oleh ahli bahasa Indonesia?
2. Bagaimana pemodelan analisis sentimen terhadap opini publik mengenai kebijakan naturalisasi pemain pada tim nasional sepak bola Indonesia dengan memanfaatkan algoritma IndoBERT *Embedding* dan Convolutional Neural Network?
3. Bagaimana performa model analisis sentimen dalam mengklasifikasikan opini publik mengenai kebijakan naturalisasi pemain pada tim nasional sepak bola Indonesia menggunakan metode IndoBERT *Embedding* dan Convolutional Neural Network?

1.4 Tujuan

Berdasarkan permasalahan yang telah dijelaskan sebelumnya, berikut merupakan tujuan dari dilakukannya penelitian ini:

1. Untuk mengetahui kecenderungan sentimen masyarakat terhadap program naturalisasi tim nasional sepak bola Indonesia berdasarkan hasil distribusi kelas dari proses pelabelan oleh ahli bahasa Indonesia.
2. Untuk mengetahui proses pemodelan analisis sentimen terhadap opini publik mengenai kebijakan naturalisasi pemain pada tim nasional sepak bola Indonesia dengan menggunakan metode *IndoBERT Embedding* dan *Convolutional Neural Network*.
3. Untuk mengevaluasi kinerja model analisis sentimen dalam mengklasifikasikan opini masyarakat terkait kebijakan naturalisasi pemain tim nasional sepak bola Indonesia dengan menggunakan metode *IndoBERT Embedding* dan *Convolutional Neural Network*.

1.5 Batasan Masalah

Untuk menjaga agar penelitian yang dilakukan tetap terfokus, maka ditetapkan beberapa batasan dalam penelitian ini:

1. Pada Penelitian ini menggunakan kumpulan *tweet* berbahasa Indonesia sebagai dataset yang relevan dengan topik program naturalisasi pemain keturunan tim nasional sepak bola Indonesia, dengan kata kunci yang digunakan seperti “Naturalisasi Timnas”, “Pemain keturunan Indonesia”, “Naturalisasi PSSI”, “Pemain Diaspora”, dan “Program Naturalisasi” serta hashtag #PemainNaturalisasi, #PemainKeturunan, #PemainDiaspora, #NaturalisasiPSSI, dan #Naturalisasi.
2. Rentan waktu *tweet* yang digunakan dalam penelitian ini adalah *tweet* yang diunggah dalam periode antara 1 Januari 2024 hingga 31 Desember 2025.
3. Pada penelitian ini, akan dibuat sebuah *single page* website yang memungkinkan pengguna untuk mengupload dataset dan menampilkan hasil analisis sentimen secara langsung dalam bentuk diagram pada halaman yang sama. Website ini bersifat sederhana dan hanya fokus pada fungsi upload serta visualisasi hasil, tanpa dilakukan pengujian lanjutan.

1.6 Manfaat Penelitian

Mengacu pada rumusan masalah serta tujuan penelitian yang telah diuraikan sebelumnya, penelitian ini diharapkan memberikan beberapa manfaat sebagai berikut:

1. Manfaat Teoritis
 - a. Memberikan kontribusi dalam pengembangan metode analisis sentimen menggunakan kombinasi IndoBERT *Embedding* dan *Convolutional Neural Network*.
 - b. Menjadi referensi teoritis bagi penelitian selanjutnya yang membahas penerapan model *deep learning* dalam pemrosesan bahasa alami.
2. Manfaat Praktis

Bagi Masyarakat dan Pengguna Media Sosial.

- a. Memberikan gambaran umum mengenai kecenderungan opini publik terhadap program naturalisasi tim nasional sepak bola Indonesia, sehingga masyarakat dapat memperoleh pemahaman yang lebih komprehensif mengenai dinamika pro dan kontra yang berkembang di media sosial.

Bagi Pengambil Kebijakan Sepak Bola Nasional

- a. Hasil penelitian ini dapat dimanfaatkan oleh pemangku kepentingan sepak bola nasional, seperti federasi atau klub sepak bola Indonesia, sebagai bahan evaluasi dalam memahami persepsi dan respons publik terhadap program naturalisasi pemain tim nasional.

Bagi Peneliti

- a. Dapat menerapkan wawasan yang didapatkan selama menempuh studi sebagai mahasiswa program studi Sistem Informasi di Universitas Pendidikan Ganesha.
- b. Mampu meningkatkan pemahaman dan pengetahuan dalam menghadapi tantangan di bidang *text mining*, khususnya analisis sentimen, dengan memanfaatkan metode IndoBERT *Embedding* dan CNN.