

BAB I

PENDAHULUAN

1.1. Latar Belakang

Sepak bola merupakan olahraga yang paling diminati di Indonesia dan dunia. Sepak bola juga menjadi sarana perjuangan nasionalisme yang dipelopori oleh Ir. Soeratin Sosrosoegondo melalui pendirian PSSI. Hingga saat ini, sepak bola tetap menjadi bagian penting dari kehidupan masyarakat Indonesia, dengan sekitar 77% penduduk menunjukkan minat terhadap olahraga ini [1]. Bahkan, enam klub sepak bola Indonesia tercatat masuk dalam daftar 200 klub terpopuler dunia. Pada era digital saat ini, media sosial telah menjadi wadah utama dalam penyebaran informasi dan pembentukan opini publik. Fenomena ini tidak hanya terjadi di ranah komunikasi pribadi, melainkan juga menyentuh berbagai sektor, termasuk dunia olahraga. Dalam konteks penelitian ini, Media Sosial X dipilih sebagai objek utama untuk menganalisis opini masyarakat terhadap performa tim nasional sepak bola Indonesia di masa kepemimpinan Shin Tae-yong. Media Sosial X merupakan platform interaktif di mana jutaan pengguna secara aktif mengekspresikan perasaan, kritik, dan dukungan melalui komentar, postingan, serta reaksi berupa emoji dan gambar. Dengan tersedianya API (*Application Programming Interface*), pengumpulan data dapat dilakukan secara otomatis dan *real-time*, sehingga memungkinkan peneliti untuk memperoleh kumpulan data yang representatif dan terus-menerus terkait dinamika opini publik.

Sistem kerja Media Sosial X yang mengintegrasikan algoritma rekomendasi dan filter informasi memungkinkan penyebaran konten secara cepat kepada penggunanya. Setiap informasi yang diunggah oleh pengguna memiliki potensi untuk menyebar luas dan memicu diskursus yang intens [2]. Data yang terkumpul tidak hanya berupa teks, tetapi juga mengandung elemen visual seperti emoji dan gambar yang memberikan gambaran nuansa emosi. Di sinilah letak keunggulan dan tantangan penggunaan Media Sosial X sebagai sumber data. Di satu sisi, data yang bersifat *real-time* dan beragam memberikan kesempatan bagi peneliti untuk menangkap opini publik secara langsung dan mendalam. Di sisi lain, variasi bahasa yang digunakan, seperti penggunaan bahasa gaul, singkatan, dan istilah lokal, serta

ketidakberaturan struktur data (*unstructured data*) menyulitkan proses analisis secara konvensional. Selain itu, fenomena penyebaran spam dan informasi palsu (hoax) turut menambah kompleksitas dalam penyaringan data yang valid dan relevan. Ketidakseimbangan distribusi sentiment di mana data yang bersifat netral sering kali mendominasi, juga dapat mengakibatkan bias dalam model klasifikasi jika tidak ditangani dengan metodologi yang tepat. Kondisi ini semakin diperparah oleh dinamika konteks sosial dan budaya yang mempengaruhi cara ekspresi emosi masyarakat di dunia maya [3].

Berdasarkan kondisi tersebut, penelitian ini berfokus pada permasalahan utama, yaitu bagaimana mengklasifikasikan sentimen dari data Media Sosial X secara akurat terkait dengan performa tim nasional sepak bola Indonesia pada era kepemimpinan Shin Tae-yong. Isu ini menjadi sangat penting mengingat opini publik yang terekam dalam media sosial dapat memberikan gambaran yang mendalam mengenai persepsi masyarakat terhadap kinerja tim nasional. Sebagai ilustrasi, saat tim nasional mencetak gol, banyak komentar positif seperti “Mantap sekali penampilan Timnas!” atau “Syukurlah, kita kembali bersinar” bermunculan, sedangkan saat terjadi kekalahan atau penampilan yang mengecewakan, reaksi negatif seperti “Timnas bermain tidak teratur” atau “Pelatih harus segera evaluasi strategi” pun tak jarang muncul. Kasus semacam ini menunjukkan bahwa data yang ada mengandung nuansa sentimen campuran, yang apabila tidak dikelola dengan metode analisis yang tepat, dapat menghasilkan interpretasi yang bias atau tidak akurat. Oleh karena itu, perumusan model analisis sentimen yang mampu menangkap keragaman ekspresi dalam data media sosial menjadi krusial dalam mendukung pengambilan keputusan strategis di bidang olahraga [4].

Berbagai pendekatan telah dikemukakan dalam literatur untuk mengatasi permasalahan analisis sentimen. Salah satu pendekatan yang sering digunakan adalah metode *lexicon-based*, yang mengandalkan kamus kata dengan nilai polaritas sentimen, seperti VADER atau SentiWordNet, untuk menilai apakah suatu teks bernuansa positif, negatif, atau netral. Meskipun metode ini memiliki keunggulan dalam kesederhanaannya, keterbatasannya terletak pada ketidakmampuannya menangani konteks kalimat yang kompleks, ironi, dan variasi

bahasa yang tidak terdapat dalam kamus yang sudah ada. Di samping itu, pendekatan *machine learning* tradisional, seperti *Support Vector Machine* (SVM), *Logistic Regression*, dan *Decision Tree* telah diterapkan dalam analisis teks dengan mengandalkan ekstraksi fitur yang teliti. Metode ini menawarkan performa yang lebih baik dibandingkan dengan pendekatan *lexicon-based* apabila didukung oleh dataset yang besar dan representatif, namun memerlukan proses pelatihan dan tahap *preprocessing* yang kompleks. Perkembangan teknologi juga telah mendorong pemanfaatan model deep learning, misalnya *Long Short-Term Memory* (LSTM) dan *Convolutional Neural Network* (CNN), yang memiliki kemampuan untuk menangkap konteks dan dependensi jangka panjang dalam data teks. Meskipun demikian, pendekatan *deep learning* memerlukan sumber daya komputasi yang besar serta waktu pelatihan yang cukup lama, sehingga tidak selalu praktis untuk penelitian dengan keterbatasan sumber daya. Tak hanya itu, terdapat pula pendekatan *hybrid* yang mencoba menggabungkan kelebihan metode *lexicon-based* dan *machine learning* untuk mencapai hasil yang optimal, namun integrasi kedua metode tersebut menambah kompleksitas dalam implementasinya.

Setiap pendekatan yang telah ada memiliki keunggulan dan kelemahan masing-masing. Metode *lexicon-based*, misalnya, menawarkan kemudahan dalam implementasi dan tidak memerlukan proses pelatihan yang intensif. Akan tetapi, metode ini sering kali gagal menangkap sentimen yang bersifat campuran atau nuansa emosional yang mendalam, terutama apabila teks mengandung bahasa gaul atau struktur kalimat yang tidak baku. Pendekatan *machine learning* tradisional dapat memberikan performa klasifikasi yang lebih baik, namun sangat bergantung pada ketersediaan dataset yang cukup besar dan *proses preprocessing* yang kompleks, yang pada gilirannya dapat meningkatkan waktu komputasi dan beban kerja. Sementara itu, meskipun model *deep learning* mampu memberikan akurasi yang tinggi dalam menangkap konteks bahasa, keterbatasan sumber daya dan waktu pelatihan membuatnya kurang ideal untuk aplikasi real-time. Pendekatan *hybrid*, walaupun menawarkan potensi integrasi kelebihan dari kedua pendekatan tersebut, sering kali menghadapi tantangan dalam penyelarasan dan optimasi sehingga tidak selalu menghasilkan solusi yang efisien.

Dengan berbagai keterbatasan tersebut, penelitian ini mengusulkan

penerapan algoritma Naïve Bayes sebagai solusi dalam analisis sentimen. Algoritma Naïve Bayes dipilih karena memiliki beberapa keunggulan yang sangat relevan dengan konteks penelitian ini. Pertama, algoritma ini mengedepankan kesederhanaan dengan asumsi bahwa setiap fitur dalam data bersifat independen satu sama lain. Asumsi sederhana ini memungkinkan perhitungan probabilitas dilakukan dengan cepat dan efisien, terutama pada dataset berukuran besar yang dihasilkan oleh Media Sosial X. Kedua, literatur sebelumnya telah menunjukkan bahwa Naïve Bayes cukup efektif untuk tugas klasifikasi teks dan analisis sentimen, asalkan didukung dengan tahap *preprocessing* yang tepat. Dengan pengembangan kamus lokal serta penyesuaian teknik tokenisasi dan normalisasi teks, algoritma Naïve Bayes dapat dioptimalkan untuk mengatasi variasi bahasa, termasuk penggunaan bahasa gaul yang umum ditemui di media sosial Indonesia. Selain itu, keunggulan lain dari Naïve Bayes terletak pada kemudahan implementasinya dan kebutuhan sumber daya komputasi yang relatif rendah jika dibandingkan dengan model deep learning. Hal ini memungkinkan penelitian untuk fokus pada validasi hasil analisis dan interpretasi data tanpa harus mengeluarkan biaya komputasi yang tinggi [5].

Dengan penerapan algoritma Naïve Bayes, model yang dikembangkan mampu mengklasifikasikan sentimen dari data Media Sosial X dengan akurasi minimal mencapai 80%. Model ini dapat secara efektif membedakan antara sentimen positif, negatif, dan netral dalam waktu pemrosesan yang singkat, sehingga memungkinkan analisis secara near real-time terhadap opini masyarakat. Hasil yang diperoleh tidak hanya memberikan gambaran mengenai persepsi publik terhadap performa tim nasional sepak bola, tetapi juga mampu mengungkap korelasi antara opini publik dan kinerja tim. Misalnya, temuan yang menunjukkan bahwa peningkatan sentimen positif sebelum dan sesudah pertandingan berkorelasi dengan peningkatan performa tim dapat menjadi indikator strategis yang berguna bagi pengelola tim nasional. Selain itu, model analisis sentimen ini dapat menyediakan insight strategis yang berguna untuk evaluasi kebijakan, seperti mendeteksi lonjakan komentar negatif yang mungkin menandakan adanya masalah dalam strategi pelatihan atau manajemen tim.

Secara keseluruhan, penelitian ini memiliki signifikansi yang tinggi dalam

konteks pengembangan teknologi informasi dan analisis data untuk bidang olahraga. Dengan mengintegrasikan pendekatan analisis sentimen melalui algoritma Naïve Bayes, penelitian ini tidak hanya berkontribusi terhadap pengembangan metode klasifikasi teks yang efisien, tetapi juga memberikan dasar bagi pengambilan keputusan strategis yang responsif terhadap dinamika opini publik. Hasil penelitian dapat memberikan dampak praktis bagi pengelola tim nasional sepak bola Indonesia dengan menyediakan informasi yang objektif dan terukur mengenai dukungan atau kritik yang muncul di ranah media sosial. Selain itu, penelitian ini juga berpotensi menjadi referensi penting bagi peneliti selanjutnya yang ingin mengkaji lebih dalam tentang pengaruh opini publik terhadap performa atlet dan manajemen tim, serta bagaimana teknologi informasi dapat dimanfaatkan untuk meningkatkan strategi kompetitif di era digital.

Penerapan algoritma Naïve Bayes sebagai metode analisis sentimen dalam penelitian ini merupakan pilihan yang tepat mengingat kesederhanaannya, efisiensi pemrosesan, dan kemampuan adaptasinya terhadap karakteristik data dari Media Sosial X. Implementasi model ini mampu mengatasi berbagai tantangan yang ada, seperti variasi bahasa dan data yang tidak terstruktur, serta memberikan hasil yang dapat dijadikan acuan dalam pengambilan keputusan strategis di dunia olahraga. Penelitian ini juga dapat membuka peluang untuk pengembangan lebih lanjut, baik dalam integrasi dengan parameter teknis dan psikologis performa tim maupun dalam pengembangan model analisis sentimen yang lebih kompleks guna menangani dinamika komunikasi di era digital.

1.2. Perumusan Masalah

1.2.1 Permasalahan (*problem statement*)

Dalam era digital, media sosial seperti Twitter menjadi platform utama bagi masyarakat untuk menyampaikan opini mereka, termasuk terhadap Tim Nasional Sepak Bola Indonesia. Namun, tingginya volume data yang tidak terstruktur, variasi bahasa, serta gaya penulisan yang berbeda menjadi tantangan dalam menganalisis sentimen secara akurat. Oleh karena itu, diperlukan suatu metode berbasis informatika yang mampu mengklasifikasikan sentimen opini masyarakat secara efektif dan efisien.

1.2.2 Metode Usulan

Penelitian ini mengusulkan penggunaan algoritma Naïve Bayes sebagai metode klasifikasi sentimen karena keunggulannya dalam menangani data teks dengan jumlah besar serta kemampuannya memberikan hasil klasifikasi yang cukup akurat. Proses analisis sentimen dilakukan melalui beberapa tahapan, yaitu pengumpulan dan preprocessing data, ekstraksi fitur menggunakan TF-IDF, klasifikasi sentimen dengan Naïve Bayes, serta evaluasi performa model menggunakan metrik akurasi, precision, recall, dan confusion matrix. Dengan metode ini, penelitian dapat menghasilkan model yang mampu mengklasifikasikan sentimen publik terhadap Tim Nasional Sepak Bola Indonesia secara akurat dan efisien.

1.2.3 Pertanyaan Penelitian

1. Seberapa akurat penerapan algoritma Naïve Bayes dalam mengklasifikasikan sentimen tweet masyarakat terhadap performa Timnas sepak bola Indonesia pada era kepelatihan Shin Tae-yong di platform X?
2. Bagaimana distribusi sentimen positif dan negatif terhadap Tim Nasional Sepak Bola Indonesia di media sosial X pada era kepelatihan Shin Tae-yong?

1.3. Tujuan & Manfaat

1.3.1 Tujuan Penelitian

1. Mengukur seberapa akurat algoritma Naïve Bayes dalam mengklasifikasikan sentimen tweet masyarakat terkait performa Timnas sepak bola Indonesia pada era kepelatihan Shin Tae-yong.
2. Menganalisis distribusi sentimen positif dan negatif terhadap Tim Nasional Sepak Bola Indonesia di media sosial X pada era kepelatihan Shin Tae-yong.

1.3.2 Manfaat:

1. Secara teoretis, penelitian ini dapat memperkaya literatur mengenai penerapan algoritma Naïve Bayes dalam analisis sentimen serta menjadi referensi bagi penelitian selanjutnya yang mengkaji pengolahan data teks dari media sosial.
2. Secara praktis, hasil penelitian ini memberikan informasi objektif yang dapat dimanfaatkan oleh pengelola timnas sepak bola Indonesia dalam mengevaluasi dan meningkatkan strategi pelatihan serta manajemen tim berdasarkan persepsi dan opini masyarakat.

1.4 Batasan-batasan

Batasan masalah dalam penelitian ini disusun untuk memastikan analisis yang dilakukan lebih terfokus dan sesuai dengan tujuan yang telah ditetapkan.

1. Data diambil dari platform media social X dan hanya tweet yang berbahasa Indonesia yang digunakan.
2. Data yang diambil hanya mencakup tweet yang berhubungan dengan sentiment terhadap timnas sepak bola Indonesia pada masa kepelatihan Shin Tae-yong
3. Hasil analisis sentiment hanya diklasifikasikan ke dalam dua kategori utama yaitu sentiment positif dan sentiment negative.
4. Kata kunci spesifik yang digunakan “Shin Tae-yong”, “STY”, “Timnas”, dan “Dendy”, yang secara umum mewakili diskusi dan perbincangan terkait timnas di media sosial X.

1.5 Sistematika Proposal

Sistematika penulisan proporsal ini dibagi menjadi 5 bab yaitu:

BAB I PENDAHULUAN

Bab I membahas mengenai latar belakang penelitian, rumusan masalah, tujuan penelitian beserta manfaat penelitian, Batasan masalah, dan sistematika penulisan skripsi.

BAB II KAJIAN PUSTAKA

5. Bab II membahas mengenai teori – teori yang berhubungan dengan analisis

sentiment terhadap timnas sepak bola Indonesia pada masa kepelatihan Shin Tae-yong menggunakan algoritma Naïve Bayes pada media social X.

BAB III METODOLOGI PENELITIAN

Bab III membahas mengenai metode penelitian yang akan digunakan, yaitu metode pengumpulan data dan metode untuk analisis data, yaitu algoritma Naïve Bayes.

BAB IV HASIL DAN PEMBAHASAN

Bab IV membahas mengenai hasil penelitian yang telah dilakukan Analisa data yang telah diperoleh

BAB V PENUTUP

Bab V membahas mengenai kesimpulan penelitian dan saran bagi pengembangan penelitian selanjutnya.

REFERENSI