BAB I

PENDAHULUAN

1.1 Latar Belakang

Melalui turnamen lokal, nasional, dan internasional, olahraga adalah salah satu hal yang secara signifikan berdampak pada keberadaan manusia dan berfungsi untuk meningkatkan reputasi bangsa dan wilayah. Negara-negara dari seluruh dunia berusaha keras untuk meraih kesuksesan dalam bidang olahraga karena pencapaian tersebut meningkatkan reputasi suatu wilayah di dunia internasional. (Riefky & Pramesti, 2020). Sepak bola adalah salah satu dari beberapa olahraga yang telah berkembang atau dimulai di banyak negara. Saat ini, hampir semua orang di dunia menyukai sepak bola, menjadikannya olahraga yang populer. Selain itu, orang-orang dari berbagai usia dan latar belakang sosial dapat menghargainya. Tidak ada olahraga yang sampai saat ini dipeluk oleh orang-orang di banyak negara di seluruh dunia dengan semangat dan gairah seperti itu. (Jon, 2022).

Ultras Garuda Indonesia adalah komunitas suporter yang mendukung Tim Nasional Sepak Bola Indonesia. Setiap kali Tim Nasional berlaga, Ultras Garuda Indonesia selalu hadir langsung di Stadion, terutama ketika pertandingan berlangsung di Indonesia. Dukungan Ultras Garuda Indonesia tidak hanya terlihat langsung di Stadion, tetapi juga memberikan semangat serta dukungan melalui media sosial. Ultras Garuda Indonesia didirikan pada 25 Maret 2015. Pada awal kemunculan Ultras Garuda Indonesia, para pendukung Tim Nasional Sepakbola Indonesia menyadari bahwa tidak ada suporter sejati yang senantiasa mendukung setiap pertandingan mereka, sehingga Ultras Garuda Indonesia terbentuk. Anggota Ultras Garuda Indonesia adalah para pendukung Indonesia yang berasal dari berbagai daerah di tanah air, yang bersatu dalam mendukung tim nasional sepak bola Indonesia (YPN Karangturi, 2021).

Timnas Indonesia memanfaatkan Instagram @timnasindonesia sebagai saluran informasi, memiliki sekitar 6 juta pengikut dan lebih dari 2 ribu unggahan. Akun ini juga telah terverifikasi oleh Instagram, yang dapat dilihat pada akun resmi Timnas yaitu @timnasindonesia, menegaskan bahwa akun Instagram Timnas Indonesia adalah

@timnasindonesia (Fadillah et al., 2023). Klub Sepak Bola Timnas Indonesia menyampaikan dan menerbitkan semua aktivitasnya melalui situs web (<https://www.pssi.org/news>).

Penggunaan media sosial Instagram oleh Timnas ini dimanfaatkan sebagai saluran informasi untuk memperkenalkan diri dan memberikan informasi kepada pengikut di akun Instagram @timnasindonesia (Utami et al., n.d.). Pemakaian media sosial dalam dunia sepak bola dapat menimbulkan sikap positif maupun negatif yang berdampak pada klub, kritik terhadap kinerja pemain, dukungan optimisme untuk pertandingan mendatang, humor, sarkasme, serta perdebatan antara suporter, nyanyian berunsur rasial, saling ejek, dan keputusan wasit yang dianggap tidak adil. Dari komentar ini, akan ditentukan sentimen yang bersifat positif, negatif, atau netral.

Secara umum, metode analisis sentimen sering digunakan, terutama dalam klasifikasi, seperti *support vector machine* dan *neural network*. Kata-kata sarkastik pada akun media sosial Twitter (X) diteliti dalam penelitian ini, “Algoritma *Multiclass SVM* untuk Teks Sarkastik di Twitter”. Hasil penelitian ini menunjukkan bahwa tingkat akurasi yang diperoleh memiliki nilai yang setara, baik melalui metode acak maupun tidak acak, yaitu 60,82% untuk metode acak dan 60,93% untuk metode tidak acak (Alita, 2021). Penelitian lain yang menggunakan model CNN menunjukkan kinerja yang luar biasa dalam klasifikasi sentimen layanan *e-commerce* di Kota Batam. Rata-rata akurasi sebesar 88,50% mengindikasikan bahwa model ini mampu mengenali sentimen secara akurat. Selain itu, nilai presisi dan recall yang tinggi pada seluruh kategori sentimen membuktikan bahwa model tidak hanya mampu mendeteksi sentimen secara tepat, tetapi juga efektif dalam meminimalkan kesalahan klasifikasi (Ridwan & Muzakir, 2022).

Pendekatan *Hybrid CNN-SVM* memadukan keunggulan *Convolutional Neural Network* (CNN) dalam ekstraksi fitur dan *Support Vector Machine* (SVM) sebagai pengklasifikasi untuk meningkatkan akurasi dalam analisis sentimen. CNN diterapkan untuk mengekstrak fitur dari teks yang telah disajikan dalam format *vektor embedding*. Arsitektur CNN dalam proses ini meliputi *Embedding Layer* yang mengubah teks menjadi vektor numerik, *Convolutional Layer* dengan *filter* untuk mengenali pola dalam teks, *Activation Function* (ReLU) untuk meningkatkan kinerja jaringan, *Pooling Layer* untuk mengurangi dimensi data sambil menjaga fitur penting, serta *Flatten Layer* yang

mengubah hasil ekstraksi fitur menjadi vektor satu dimensi. Fitur yang telah diambil oleh CNN selanjutnya digunakan sebagai masukan untuk SVM (*Support Vector Machine*) untuk mengkategorikan sentimen. SVM memanfaatkan kernel seperti Linear Kernel untuk data linier atau RBF Kernel (*Radial Basis Function*) untuk data *non-linier*, dengan tujuan mencari *hyperplane* terbaik yang mengelompokkan data ke dalam kategori emosi positif dan negatif. Setelah pelatihan model, evaluasi dilakukan dengan memanfaatkan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*, serta matriks kebingungan untuk mengamati jumlah prediksi yang akurat dan keliru. Pendekatan *Hybrid CNN-SVM* ini memiliki sejumlah kelebihan, seperti menyatukan keunggulan CNN dalam ekstraksi fitur dan kemampuan SVM dalam klasifikasi, menurunkan risiko *overfitting* yang sering muncul pada penggunaan CNN saja, serta memberikan performa yang lebih unggul dibandingkan dengan model yang hanya menerapkan CNN atau SVM secara terpisah.

Untuk proyek akhir ini, penelitian yang akan dipaparkan adalah “Analisis Sentimen Klub Tim Nasional Menggunakan *Hybrid CNN-SVM*”, yang didasarkan pada berbagai penjelasan latar belakang yang telah diberikan.

1.2 Perumusan Masalah

Bagaimana menggunakan pendekatan *Hybrid CNN-SVM* untuk analisis sentimen pada akun media sosial Instagram Tim Sepak Bola Nasional adalah pertanyaan penelitian yang dapat dikembangkan berdasarkan kesulitan-kesulitan yang telah disebutkan?

1.3 Batasan Masalah

Penentuan ruang lingkup penelitian bertujuan untuk meningkatkan ketepatan dan fokus penelitian. Batasan ruang lingkup yang diterapkan dalam penelitian ini adalah sebagai berikut:

- a. Unggahan media sosial Instagram klub sepak bola Tim Nasional menjadi sumber data penelitian ini, yang dikumpulkan antara 18 Agustus 2023 hingga 16 November 2024.
- b. Penelitian ini difokuskan pada analisis komentar teks yang berasal dari 90 pertandingan Tim Nasional, yang terdiri atas 30 postingan saat Tim Nasional

meraih kemenangan, 30 postingan saat mengalami kekalahan, dan 30 postingan saat pertandingan berakhir seri.

1.4 Tujuan dan Manfaat

Hasil penerapan metode *Hybrid CNN-SVM* dalam penelitian ini menghasilkan klasifikasi data komentar pada akun Instagram Timnas dibagi menjadi tiga kategori sentimen: netral, negatif, dan positif. Penelitian masa depan tentang analisis media sosial Instagram dengan topik terkait diharapkan dapat menggunakan studi ini sebagai referensi. Selain itu, studi ini juga bertujuan memberikan wawasan mengenai penerapan metode *Hybrid CNN-SVM* dalam analisis sentimen pada platform media sosial, khususnya di bidang sepak bola maupun topik-topik lain yang relevan.

1.5 Metode Penelitian

Proses metodologi penelitian terhadap komentar akun Instagram Timnas Indonesia dilakukan melalui beberapa tahapan. Berikut adalah langkah-langkah yang akan dilaksanakan dalam penelitian ini:

a. *Data Selection*

Tahap Memilih data yang relevan dan tepat untuk digunakan dalam proses klasifikasi adalah langkah pertama dalam proses *data mining*.

b. *Data Cleaning* (Pre-processing)

Pada tahap ini, pra-pemrosesan data dilakukan, yang melibatkan pembersihan data penelitian agar lebih terorganisir, bersih, dan siap untuk dianalisis. Untuk membersihkan data, karakteristik atau informasi yang tidak diperlukan untuk proses kategorisasi dan prediksi dihilangkan.

c. *Data Labelling* (Pelabelan Data)

Data sekarang akan diklasifikasikan sebagai positif, negatif, atau netral berdasarkan kategori sentimen setelah langkah pra-pemrosesan.

d. Klasifikasi

Pada tahap ini, dilakukan penerapan teknik *data mining* untuk memproses data, dengan *output* berupa nilai akurasi, presisi, dan *recall*.

e. *Data Testing* (Pengujian Data)

Prosedur pengujian data digunakan pada titik ini untuk mengonfirmasi bahwa setiap ulasan secara akurat termasuk dalam kategori sentimen positif, negatif, atau netral.

1.6 Sistematika Penulisan

Untuk membantu pembaca memahami isi laporan ini, laporan akhir proyek ini dibagi menjadi lima bab, masing-masing berisi pembahasan yang sistematis dan terstruktur. Format laporan ini sebagai berikut:

BAB I PENDAHULUAN

Latar belakang, rumusan masalah, batasan penelitian, tujuan, manfaat, metodologi penelitian, dan sistem penulisan semuanya dijelaskan dalam bab ini.

BAB II TINJAUAN PUSTAKA

Selain metode TF-IDF, teori *Convolutional Neural Networks* (CNN), teori *Support Vector Machines* (SVM), konsep *Hybrid CNN-SVM* yang digunakan dalam studi ini, serta ide dasar analisis sentimen dan metode analisis sentimen, bab ini juga membahas studi teoretis dan penelitian sebelumnya yang berkaitan dengan topik-topik yang dibahas dalam bab ini.

BAB III METODOLOGI PENELITIAN

Kerangka penelitian, fase pengumpulan data beserta penjadwalan dan prosedurnya, pemilihan populasi dan sampel, pra-pengolahan data, pendekatan analisis, visualisasi hasil, dan skenario pengujian semuanya dibahas dalam bab ini.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi penyajian hasil penelitian, mulai dari spesifikasi perangkat, hasil pengumpulan dan pra-pemrosesan data, pelabelan data, pembentukan model *Hybrid CNN-SVM*, evaluasi performa dengan *confusion matrix*, hingga analisis hasil klasifikasi sentimen.

BAB V KESIMPULAN DAN SARAN

Termasuk kesimpulan yang dihasilkan dari temuan studi, masalah yang dihadapi, dan rekomendasi untuk penelitian lebih lanjut.