

BAB I

PENDAHULUAN

1.1 Latar Belakang

Konflik adalah tindakan atau peristiwa yang sangat merugikan karena dapat menyebabkan kerusakan, baik terhadap pihak yang terlibat maupun orang lain, termasuk kehilangan nyawa dan harta benda. Semua bentuk konflik harus segera dihentikan, termasuk konflik antara Palestina dan Israel. Palestina dan Israel adalah dua negara di Timur Tengah yang masih mengalami konflik hingga saat ini[1]. Sejarah konflik Palestina dimulai sejak masa pemerintahan Bani Israil hingga era Nabi Daud, yang menjadi penguasa Bani Israil, kemudian dilanjutkan oleh putranya, Nabi Sulaiman. Setelah kematian Nabi Sulaiman, kerajaan tersebut terpecah menjadi dua: Kerajaan Yehuda dan Kerajaan Israel. Konflik yang berkelanjutan dimanfaatkan oleh Kekaisaran Asyur yang menaklukkan Kerajaan Israel pada tahun 722 SM, menggulingkan banyak penduduknya, dan menyerang Yerusalem[2].

Menurut data dari Kementerian Komunikasi dan Informatika, terdapat sekitar 63 juta pengguna internet di Indonesia. Angka ini menunjukkan bahwa masyarakat Indonesia sangat aktif dalam menggunakan internet, termasuk media sosial. Di seluruh dunia, termasuk Indonesia, orang-orang memanfaatkan media sosial untuk mencari informasi, hiburan, relaksasi, serta untuk berinteraksi secara sosial[3]. Teknologi media sosial telah mengubah cara orang berkomunikasi, menyebarkan informasi, membangun hubungan, dan menjadi bagian penting dalam kehidupan sehari-hari[4]. Media sosial memiliki kekuatan yang signifikan dalam memengaruhi opini publik yang berkembang di masyarakat, sehingga menghasilkan pengetahuan baru[5].

Pesatnya perkembangan media sosial dan platform daring lainnya telah menghasilkan banyak data, termasuk opini dan sentimen dari masyarakat. Media sosial yang digunakan dalam penelitian ini adalah X, platform yang populer dan sering digunakan sebagai sumber data untuk analisis tren serta opini publik[6]. Konflik antara Israel dan Palestina menarik perhatian dunia. Isu ini menjadi perbincangan global, termasuk di Indonesia, yang memunculkan berbagai opini publik, terutama di media sosial X.

Penelitian ini bertujuan untuk membandingkan metode klasifikasi dalam mengidentifikasi sentimen publik dari data tweet pengguna X. Subjek penelitian ini adalah data tweet yang memuat tagar #israel dan #palestina[7]. Informasi terkait konflik Palestina-Israel tersebar luas di media sosial, mendorong masyarakat untuk memberikan opini dan sentimen terhadap berita yang mereka terima.

Pandangan dan respons masyarakat sangat bernilai. Melalui tweet yang dipublikasikan, para peneliti dapat mengetahui perspektif pengguna dengan menganalisis setiap tweet. Ini dapat memperdalam jurang perpecahan dalam masyarakat, terutama karena kebencian terhadap salah satu pihak. Kondisi ini muncul karena masalah sensitif yang sering dikaitkan dengan konflik atau peperangan. Namun, karena jumlah data yang sangat besar, diperlukan suatu cara untuk mengubahnya menjadi informasi yang relevan. Untuk mencapai hal ini, *sentiment analysis* digunakan. *Sentiment analysis* menganalisis teks digital untuk menentukan apakah pesan tersebut bersifat positif atau negatif[8].

Sentiment analysis mencakup beberapa bidang seperti *text mining*, *natural language processing* (NLP), dan *artificial intelligence*, yang digunakan untuk secara otomatis mengekstraksi, memahami, dan memproses data teks guna mendapatkan informasi serta pengetahuan baru. Proses *sentiment analysis* menentukan bagaimana masyarakat meninjau suatu ulasan terhadap isu-isu yang beredar dan cenderung membagi sentimen menjadi dua atau lebih kategori[9]. Penelitian ini akan membandingkan dua algoritma untuk menganalisis sentimen mengenai konflik Palestina-Israel, yaitu algoritma Random Forest dan BERT. Tujuan penelitian ini adalah untuk mengkategorikan opini masyarakat melalui media sosial X menjadi opini positif dan negatif.

Random Forest adalah metode yang menggunakan teknik *ensemble learning*, di mana beberapa model pohon keputusan digabungkan untuk mencapai hasil klasifikasi yang lebih akurat dan stabil[10]. Selama proses pelatihan, Random Forest mengambil sampel data pelatihan secara acak. Algoritma ini dapat mengklasifikasikan sejumlah besar data dengan akurat, dengan hasil akhir diperoleh melalui penentuan node akar dan berakhir pada beberapa node daun[11]. Random Forest dipilih karena kemampuannya yang sangat baik dalam mengelola data yang kompleks dan beragam, serta menghasilkan hasil yang akurat untuk klasifikasi dan prediksi. Metode ini ideal untuk *sentiment analysis* karena mampu

mengatasi *overfitting*, efektif pada dataset besar dan tidak terstruktur, serta memberikan pemahaman yang jelas tentang fitur-fitur yang memengaruhi hasil[12].

BERT adalah model pembelajaran mendalam yang memberikan hasil canggih dalam berbagai tugas Pemrosesan Bahasa Alami (*Natural Language Processing* atau NLP). BERT mampu menghasilkan model bahasa dua arah yang mendalam dengan memahami konteks dari teks. Dalam BERT, terdapat enam lapisan *transformer* di setiap *encoder* dan *decoder*. Google telah membuka akses ke model BERT yang sebelumnya terkunci, yang memungkinkan penggunaan model ini dimulai dari representasi kata dengan *word embeddings*, salah satu keuntungan utama dari BERT[13]. Dalam pengembangan model BERT pada penelitian ini, digunakan metode *fine-tuning*, yang terbukti secara signifikan meningkatkan performa model[14]. Pemanfaatan model BERT (*Bidirectional Encoder Representations from Transformers*) dalam *sentiment analysis* telah terbukti sangat efektif menurut berbagai penelitian ilmiah. BERT memiliki kemampuan istimewa dalam memahami bahasa secara menyeluruh, yang memungkinkan model ini menghasilkan representasi teks yang lebih tepat dan informatif[15].

Membandingkan algoritma Random Forest dan BERT dalam *sentiment analysis* sangat penting karena keduanya memiliki keunggulan dan keterbatasan masing-masing. Random Forest adalah metode *ensemble learning* yang efisien dan mudah diinterpretasikan, tetapi sering kali kurang dalam menangkap konteks dan nuansa bahasa, yang dapat mengurangi akurasi dalam analisis sentimen yang kompleks. Sebaliknya, BERT, dengan arsitektur *Transformer*-nya, menawarkan pemahaman kontekstual yang mendalam dan akurasi yang tinggi, namun memerlukan sumber daya komputasi yang lebih besar serta lebih kompleks dalam pelatihan dan interpretasi.

Berdasarkan latar belakang tersebut, penelitian ini diberi judul “*Perbandingan Algoritma Random Forest dan BERT Untuk Analisis Sentiment Palestina-Israel Pada Sosial Media X*” Penelitian ini dirancang untuk mengevaluasi dan membandingkan efektivitas dua algoritma *sentiment analysis* dalam memahami opini publik terkait konflik Palestina-Israel di platform media sosial X. Dengan memanfaatkan data tweet yang mengandung tagar #israel dan #palestina, penelitian ini bertujuan untuk mengidentifikasi dan mengkategorikan sentimen yang berkembang di masyarakat, baik yang positif maupun

negatif. Algoritma Random Forest dan BERT akan digunakan untuk menganalisis sentimen dari data tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan diatas, penulis memiliki beberapa rumusan masalah terkait permasalahan tersebut :

1. Bagaimana penerapan algoritma BERT dan random forest dalam analisis sentiment mengenai konflik palestina-israel di media X?
2. Bagaimana perbandingan kinerja algoritma BERT dan random forest dalam analisis sentiment mengenai konflik palestina-israel di media X?

1.3 Batasan Masalah

Berdasarkan rumusan masalah diatas, peneliti membatasi masalah yang akan dianalisa pada pembuatan sistem ini, Batasan masalah tersebut antara lain :

1. Analisis sentiment menggunakan algoritma BERT dan Random forest sebagai fokus utama penelitian
2. Penelitian ini membatasi analisis terhadap periode waktu tertentu, yaitu dari bulan Juli sampai dengan bulan Oktober.
3. Fokus pada isu konflik palestina-israel dengan penekanan pada topik topik tertentu seperti konflik terkini atau isu isu terkait.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, penulis memiliki beberapa tujuan terkait dalam permasalahan tersebut, yaitu:

1. Menerapkan algoritma BERT dan random forest dalam analisis sentiment mengenai konflik palestina-israel di media X
2. Menentukan perbandingan kinerja algoritma BERT dan random forest dalam analisis sentiment mengenai konflik palestina-israel di media X.

1.5 Manfaat Penelitian

Manfaat penelitian ini dapat mencakup beberapa aspek, seperti:

1. Pemahaman publik dan pengambilan keputusan

Menyediakan pemahaman yang lebih baik tentang bagaimana public merespons Konflik palestina-israel, memberikan informasi berharga bagi pengambil keputusan dan pembuat kebijakan

2. Perkembangan algoritma analisis sentiment

Menyumbangkan kontribusi untuk pengembangan dan peningkatan algoritma analisis sentiment, terutama dalam konteks konflik palestina-israel.

3. Pengembangan strategi komunikasi

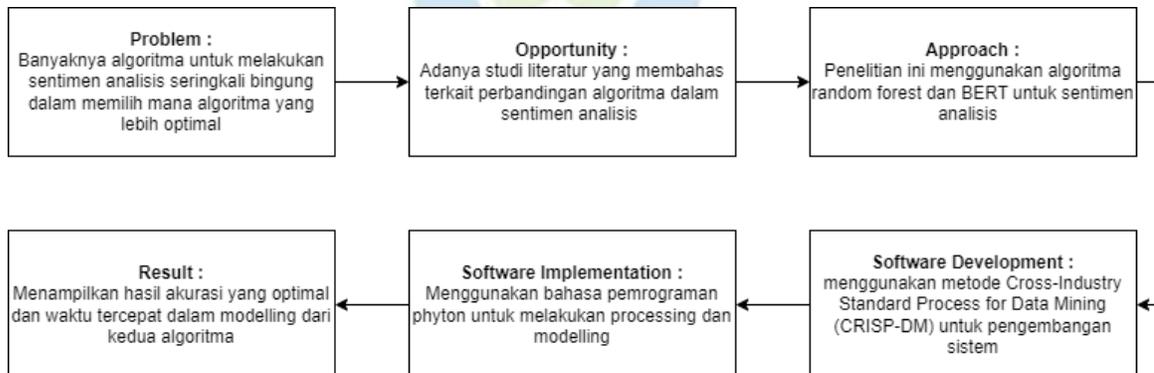
Membantu dalam merancang strategi komunikasi yang lebih efektif terkait dengan isu konflik palestina-israel, baik untuk organisasi media maupun pihak yang terlibat dalam komunikasi public

4. Kontribusi terhadap penelitian selanjutnya

Menjadi dasar untuk penelitian lebih lanjut terkait dengan analisis sentimen, pengaruh media, dan opini publik dalam konteks isu konflik palestina-israel.

1.6 Kerangka Pemikiran

Adapun untuk Kerangka pemikiran dari penelitian ini terdapat pada Gambar 1.1:



Gambar 1. 1 Kerangka Pemikiran

1.7 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini disusun menjadi lima bab, di mana setiap bab dirancang untuk memenuhi kebutuhan dalam penelitian yang membahas perbandingan algoritma Random Forest dan BERT, mulai dari latar belakang permasalahan, tujuan, landasan teori, analisis perancangan, hingga implementasi dan pengujian sistem. Adapun sistematika penulisan pada setiap bab adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang yang menjelaskan alasan permasalahan tersebut diangkat dalam penelitian, perumusan masalah, batasan masalah, tujuan, metodologi penelitian, serta sistematika penulisan yang digunakan sebagai dasar dalam perancangan dan pembuatan tugas akhir.

BAB II KAJIAN LITERATUR

Bab ini membahas landasan teori dan teori-teori pendukung yang relevan dengan perbandingan algoritma Random Forest dan BERT dalam analisis sentimen. Pembahasan mencakup teori dasar kedua algoritma, proses kerja, serta studi terkait yang mendukung penelitian ini.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode yang digunakan dalam penelitian, termasuk analisis sistem yang dirancang. Perancangan sistem mencakup langkah-langkah preprocessing data teks, pemilihan algoritma, serta pengolahan data menggunakan Random Forest dan BERT. Juga dijelaskan analisis kebutuhan, analisis data, dan evaluasi kelayakan algoritma yang dibandingkan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil dari eksperimen yang dilakukan dengan kedua algoritma, yaitu Random Forest dan BERT. Selain itu, dilakukan pengujian performa, analisis hasil, serta perbandingan antara kedua algoritma dalam konteks akurasi, efisiensi, dan kapabilitas dalam menangani analisis sentimen.

BAB V KESIMPULAN

Bab ini berisi kesimpulan dari hasil perbandingan algoritma Random Forest dan BERT, termasuk kelebihan dan kekurangan masing-masing algoritma. Selain itu, diberikan saran untuk pengembangan lebih lanjut dalam peningkatan akurasi dan efisiensi algoritma dalam analisis sentimen.