

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pesatnya perkembangan teknologi komunikasi menjadi tidak terpisahkan dalam kehidupan masyarakat. Diikuti dengan perkembangan internet, ruang lingkup masyarakat kini semakin erat hubungannya dengan media komunikasi. Sebuah riset menyatakan bahwa Indonesia merupakan salah satu negara dengan populasi pengguna internet terbesar di dunia. Data pengguna internet di Indonesia menurut *We are Social* dan *Hootsuite* yang dijelaskan pada Gambar 1.1 didapatkan hasil 74 % penduduk Indonesia terhubung oleh internet dari total populasi 277,7 juta jiwa. Pernyataan dalam riset ini juga membuktikan bahwa 68,9% atau 191 juta penduduk Indonesia aktif menggunakan media sosial[1].

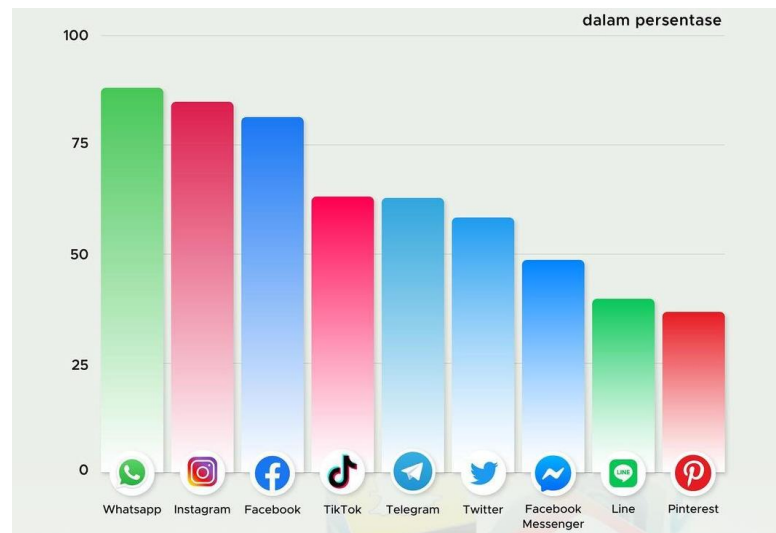


Gambar 1.1 Jumlah Pengguna Internet di Indonesia

Berkembangnya teknologi komunikasi berdampak terhadap munculnya rasa kekhawatiran akan efek media sosial pada masyarakat. Media sosial saat ini telah menjadi populer di kalangan masyarakat dalam menjalin komunikasi antara satu dengan yang lainnya. Pemanfaatan media sosial seakan menjadi jawaban atas kebutuhan komunikasi yang serba cepat, mudah, dan akurat. Di era saat ini media sosial memiliki peranan penting bukan hanya untuk komunikasi, tetapi juga dijadikan sebagai sarana dalam menyampaikan suatu aspirasi bagi masyarakat mengenai masalah kesehatan, ataupun masalah yang berhubungan dengan pemerintah, politik, dan permasalahan lainnya.

Menurut McGraw Hill Dictionary, media sosial adalah sarana yang digunakan oleh orang-orang untuk berinteraksi satu sama lain dengan cara menciptakan, berbagi, serta bertukar informasi dan gagasan dalam sebuah jaringan dan komunitas virtual. Salah satu media sosial yang paling populer dari sekian banyak *platform* di kalangan masyarakat, *Twitter* menjadi salah satu yang favorit digunakan. Berdasarkan laporan *We are Social* sepanjang tahun 2022 yang

dijelaskan pada Gambar 1.2, jumlah aktif pengguna twitter di Indonesia mencapai 18,45 juta pengguna, angka tersebut membuat Indonesia menempati peringkat kelima negara pengguna Twitter terbesar di dunia[2]. Banyak alasan mengapa pengguna memilih Twitter, salah satunya untuk melakukan komunikasi dan menyampaikan opini terhadap berbagai topik yang sedang ramai dibicarakan.



Gambar 1.2 Jumlah Pengguna Twitter di Indonesia

Twitter adalah layanan jejaring sosial *microblogging* gratis yang memungkinkan pengguna yang terdaftar menyiarkan pesan singkat yang disebut *tweets*. Data *tweet* yang diutarakan di *Twitter* dapat dijadikan sebagai parameter untuk mengukur tingkat kepuasan publik tentang berbagai topik, hal ini menjadikan *Twitter* sebagai sumber media untuk mendapatkan data tekstual yang sangat besar dan berpotensi menyimpan informasi dengan kandungan sentimen dalam satu kalimat *tweet*[3]. Studi kasus yang diambil oleh peneliti adalah tragedi di Stadion Kanjuruhan Malang. Pengambilan studi kasus Tragedi Kanjuruhan dilakukan karena kasus ini banyak menarik minat masyarakat untuk menyampaikan ide dan komentar terkait tragedi tersebut di media sosial.

Komentar atau *tweet* tersebut mewujudkan pendapat dan tanggapan yang berbeda dari setiap pengguna media sosial. Melihat perbedaan pendapat antar *suporter* di media sosial telah memunculkan kekhawatiran mengenai pro kontra penggunaan gas air mata, serta ketidak tegasan aparat dalam mengatasi kasus kanjuruhan yang menyebabkan terjadinya pelanggaran HAM, sehingga dalam perkembangan kasusnya sulit untuk membedakan secara digital maksud dari kecenderungan *tweet*. Maka dari itu perlu dilakukan analisis terhadap setiap *tweet* tersebut, analisis yang tepat digunakan untuk menganalisa sentimen masyarakat dan mengkategorikan

tweet ke dalam kelas positif, negatif, dan netral secara otomatis dengan menerapkan analisis sentimen. Analisis sentimen atau *opinion mining* dilakukan untuk memperoleh persepsi atau gambaran umum masyarakat terhadap proses keadilan Tragedi Kanjuruhan.

Analisis sentimen atau *opinion mining* adalah bidang studi yang mempelajari cara untuk menganalisa opini, sentimen, evaluasi, penilaian, sikap, dan emosi terhadap entitas seperti produk, jasa, organisasi, peristiwa, isu topik dengan berbagai atributnya. Proses ini dilakukan sebagai bentuk penerapan dari *Natural Language Processing* dalam menganalisis teks. Sentimen dapat ditemukan di dalam sebuah komentar yang memberikan indikator untuk menanggapi suatu topik, sentimen dikategorikan menjadi dua kelas yaitu positif dan negatif. Dalam hal ini analisis sentimen diimplementasikan untuk berfokus pada tugasnya dalam mengolah setiap kategori sentimen. Banyak penelitian yang menggunakan topik analisis sentimen sebagai sarana untuk mengetahui secara akurat maksud yang disampaikan pengguna pada *tweet* tersebut. Sehingga didapat informasi mengenai kualitas produk, strategi penerimaan produk, dan analisa sentimen publik yang sehubungan dengan kebijakan dan layanan publik[4].

Tahapan dalam analisis sentimen, terdapat beberapa algoritma yang banyak digunakan diantaranya adalah *Naïve Bayes*, *Support Vector Machine*, *K Nearest Neighbour*, dll. Namun seiring dengan perkembangan ilmu pengetahuan, berkembang pula metode-metode pada *opinion mining*. Muncul model yang lebih modern yaitu model *deep learning*, beberapa contoh model *deep learning* adalah *Recurrent Neural Network* (RNN), *Convolutional Neural Network* (CNN), *Long Short-Term Memory* (LSTM), dll. Pada dasarnya metode *deep learning* difokuskan untuk klasifikasi pengolahan data citra, namun terdapat dalam sebuah penelitian yang mengusulkan penggunaan kombinasi CNN dan LSTM yang diuji pada dimensi analisis sentimen. Usulan konsep baru penggunaan metode CNN-LSTM dalam pengolahan teks menunjukkan bahwa CNN dan LSTM merupakan metode yang unggul dan tepat dalam klasifikasi data teks[5].

Oleh karena itu metode yang diusulkan dalam penelitian ini adalah metode *deep learning* yang terdiri dari gabungan metode *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM). Metode gabungan tersebut merupakan metode *deep learning* yang banyak digunakan dan terbukti memiliki hasil *accuracy* yang baik dalam melakukan pemrosesan data tekstual ke dalam kelas positif, negatif, dan netral. Sehingga menggunakan metode CNN-LSTM mampu menghasilkan nilai akurasi yang tepat dan akurat. Dengan adanya kombinasi CNN-LSTM akan meningkatkan performa analisis sentimen.

Proses pada algoritma *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM) yaitu terdiri dari konvolusi awal yang akan menerima penyisipan kata sebagai masukan. Kemudian *output* akan dikumpulkan ke dimensi yang lebih kecil yang kemudian masuk ke dalam lapisan LSTM. Kunci utama dari metode ini adalah lapisan konvolusi akan mengekstraksi fitur awal dan metode LSTM akan melakukan klasifikasi fitur untuk menentukan pengurutan kategori sentimen yang masuk. Berdasarkan uraian permasalahan diatas, algoritma CNN-LSTM mampu melakukan ekstraksi yang cepat dan tepat. Sehingga hasil dari klasifikasi *tweet* yang didapatkan pada topik Tragedi Kanjuruhan dapat dijadikan sebagai bahan untuk mencari hasil klasifikasi sentimen pengguna *Twitter* berdasarkan kategori positif, negatif, dan netral dengan tingkat akurasi terbaik terhadap opini publik mengenai suatu informasi yang sedang ramai dibicarakan di media sosial.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Insiden tragis pasca pertandingan Arema FC melawan Persebaya di Kanjuruhan telah berhasil menyita perhatian publik. Ramainya pemberitaan tentang tragedi Kanjuruhan membuat *Twitter* dipenuhi cuitan atau ulasan masyarakat terhadap peristiwa tersebut, sehingga tragedi kanjuruhan menjadi *trend* topik yang banyak dibicarakan di Indonesia. Sentimen yang terkandung dalam ulasan atau opini publik dapat mempengaruhi pandangan masyarakat terhadap suatu kebijakan atau pelayanan pemerintah, sehingga masyarakat harus bisa membedakan opini positif, negatif, dan netral. Untuk itu, perlu dilakukan pembuatan sistem untuk memberikan sentimen otomatis dan mengklasifikasi sentimen agar masyarakat dapat membedakan opini positif, negatif, dan netral terhadap tragedi tersebut.

1.2.2 Metode Usulan

Dalam penelitian ini mengusulkan pengujian klasifikasi sentimen masyarakat mengenai tragedi Kanjuruhan pada data *Twitter* menggunakan kombinasi metode CNN (*Convolutional Neural Network*)- *Long Short-Term Memory* (LSTM). Metode ini akan digunakan untuk mengklasifikasikan sentimen masyarakat sesuai terhadap kelas positif, negatif, dan netral dengan data ulasan sentimen yang diambil dari *Twitter*.

1.2.3 Pertanyaan Penelitian

Berdasarkan latar belakang permasalahan diatas maka dapat dirumuskan beberapa pertanyaan pada penelitian ini yaitu :

1. Bagaimana metode CNN (*Convolutional Neural Network*) – LSTM (*Long Short-Term Memory*) mampu mengklasifikasi opini masyarakat mengenai tragedi Kanjuruhan?
2. Bagaimana hasil performa kinerja sistem pada analisis sentimen menggunakan metode CNN (*Convolutional Neural Network*) – LSTM (*Long Short-Term Memory*)?

1.3 Tujuan dan Manfaat Penelitian

1.3.1 Tujuan Penelitian

Berikut adalah tujuan penelitian :

1. Mengaplikasikan metode CNN (*Convolutional Neural Network*) – LSTM (*Long Short-Term Memory*) untuk klasifikasi ulasan mengenai studi kasus sentimen tragedi kanjuruhan Malang pada data *Twitter*.
2. Mengetahui hasil kinerja sistem pada analisis sentimen menggunakan metode CNN (*Convolutional Neural Network*) – LSTM (*Long Short-Term Memory*).

1.3.2 Manfaat Penelitian

Adapun manfaat penelitian ini adalah :

1. Mahasiswa dapat menggunakan metode CNN (*Convolutional Neural Network*) – LSTM (*Long Short-Term Memory*) untuk melakukan analisis sentimen terhadap dataset komentar di *Twitter* terhadap berbagai topik yang sedang ramai dibicarakan.
2. Menjadi bahan referensi bagi para penelitian selanjutnya yang relevan dengan kasus penelitian analisis sentimen.
3. Bagi masyarakat, bisa dijadikan acuan untuk mengetahui pandangan masyarakat mengenai kebijakan atau layanan pemerintah dalam menangani kasus tragedi Kanjuruhan.

1.4 Batasan Masalah

Adapun batasan-batasan masalah pada penelitian ini adalah :

1. Data topik yang digunakan pada penelitian ini adalah data cuitan terkait dengan Tragedi Kanjuruhan di *Twitter*.
2. Data *Twitter* yang diperoleh menggunakan Teknik *Scraping* dengan menggunakan *library snsrape* pada Python dengan pembatasan *keywords* yaitu “kasus kanjuruhan” dan “tragedi kanjuruhan”.
3. Data sentimen *Twitter* memanfaatkan tenaga ahli yaitu pengajar bahasa Indonesia untuk mendapatkan hasil ulasan sentimen positif, negatif, dan netral.
4. Algoritma atau metode yang digunakan dalam penelitian ini adalah kombinasi metode CNN (*Convolutional Neural Network*) – LSTM (*Long Short-Term Memory*).
5. Pengambilan data *Twitter* dengan periode waktu 24 Oktober-1 Desember 2022.
6. Data *Twitter* yang diambil hanya ulasan berbahasa indonesia sebanyak 4500 *tweet*.

1.5 Sistematika Penelitian

Sebagai acuan bagi penulis agar penulisan skripsi ini lebih terarah dan tersusun dengan rapi sesuai dengan yang penulis harapkan, maka penulisan skripsi ini akan disusun secara sistematis dan dibagi menjadi beberapa BAB, seperti berikut :

1. BAB I PENDAHULUAN

Bab ini berisi tentang gambaran umum dari penulisan tugas akhir yang meliputi latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, dan sistematika penulisan.

2. BAB II TINJAUAN PUSTAKA

Bab ini menguraikan tentang tinjauan pustaka yang meliputi analisis sentimen, media sosial *Twitter*, *natural language processing*, *deep learning*, *text preprocessing*, metode CNN, metode LSTM, dan metode CNN-LSTM. Selanjutnya landasan teori keilmuan yang memuat penelitian terdahulu yang menjadi referensi penelitian. Bagian ini juga berisi tentang dasar teori dan metode yang akan diimplementasikan pada topik skripsi ini.

3. BAB III METODOLOGI PENELITIAN

Pada bab ini berisi mengenai metodologi penelitian yang membahas langkah-langkah yang akan dilakukan dalam penelitian ini. Bagian ini akan menjelaskan gambaran umum penelitian, pengumpulan data, alur pemodelan yang digunakan, dan kerangka penelitian.

4. BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan lebih rinci mengenai penerapan dan hasil klasifikasi sentimen yang akan dibentuk. Langkah-langkah penerapan tersebut meliputi visualisasi data, tahap *preprocessing* data, ekstraksi fitur atau transformasi data dengan Word2vec, pembuatan model klasifikasi CNN-LSTM (*Convolutional Neural Network*) – *Long Short-Term Memory*) sebagai acuan dalam implementasi sistem, dan langkah terakhir adalah pengujian sistem yang telah diimplementasi dengan melakukan proses verifikasi dan validasi beserta pengujian kinerja dari sistem yang dibuat.

5. BAB V PENUTUP

Bab ini berisi kesimpulan yang didapatkan dari penelitian yang telah dilakukan serta saran untuk pengembangan penelitian selanjutnya.

6. DAFTAR PUSTAKA

Pada bab ini berisi mengenai daftar sumber yang telah digunakan untuk referensi sebagai sumber tambahan informasi dalam mengerjakan proposal penelitian.