

BAB I PENDAHULUAN

1.1 Latar Belakang

Abad 21 sering disebut sebagai abadinya informasi. Hal ini dikarenakan penggunaan sosial media seperti Facebook, Twitter, Pinterest, dan youtube pada abad ini mengalami kenaikan yang signifikan dibandingkan dengan tahun-tahun sebelumnya[1]. Twitter merupakan salah satu media sosial yang cukup digandrungi oleh masyarakat. Pengguna Twitter di Indonesia pada quartal 3 tahun 2017 mencapai 49% dari jumlah penduduk Indonesia, dimana mayoritas pengguna sangat variative mulai dari kalangan remaja berumur 18 sampai orang dewasa umur 60 tahun[2]. Pengguna Twitter dapat melakukan 400 lebih tweets per harinya. Para peneliti banyak melakukan penelitian analisis teks yang memanfaatkan data dari twitter. Hal ini karena Twitter memiliki banyak keunggulan, seperti jumlah karakter yang banyak (lebih dari 140 karakter), jumlah teks yang sangat banyak, format pesan yang tidak formal, latar belakang user yang variatif, dan sangat mudah untuk digunakan[1].

Twitter sebagai *microblogging* banyak digunakan oleh masyarakat untuk memberikan apresiasi dan menyatakan opini terhadap suatu hal, memposting dan memberikan rating terhadap sebuah topik, jasa, atau produk. Secara umum pemakaian Twitter banyak digunakan untuk menyatakan sebuah opini atau pendapat pribadi terhadap perihal yang spesifik hingga perihal yang umum (*public statement*)[1]. Namun data yang didapat dari Twitter merupakan data yang tidak terstruktur sehingga membutuhkan metode *Deep Learning* yang akan memproses teks dari data yang didapatkan. *Deep Learning* merupakan sebuah metode pembelajaran yang dilakukan oleh komputer yang memiliki beberapa tingkatan abstraksi dan representasi guna membantu proses memahami data seperti suara, teks, dan gambar[3]. Dimana hasil pengolahan data dari *Deep Learning* adalah untuk menganalisis sentiment. Analisis Sentimen (*Sentiment Analysis*) merupakan sebuah riset secara terkomputasi terhadap opini sentimen dan emosi yang dikemas secara tekstual[3].

Pada metode deep learning dapat diterapkan kombinasi dua algoritma untuk meningkatkan akurasi dalam proses klasifikasi terhadap data testing, salah satunya

adalah kombinasi algoritma *Convolutional Neural Network* (CNN) dan algoritma *Long Short Term Memory* (LSTM). Kombinasi dua algoritma ini sangat cocok untuk proses klasifikasi karena keunggulan yang dimiliki oleh masing-masing algoritma. Algoritma CNN memiliki keunggulan dalam mendeteksi dan mengenali pola dari suatu objek, baik itu berupa gambar ataupun teks. Algoritma LSTM sangat cocok digunakan untuk mempelajari pola dari data *sequential* karena algoritma LSTM memiliki jumlah sel yang lebih banyak dan dapat menghindari *vanish gradient*.

Telah banyak penelitian – penelitian sebelumnya yang memiliki *related* arah penelitiannya, seperti pada penelitian yang berjudul “Sentimen Analisis Tweet Berbahasa Indonesia dengan *Deep Belief Network*” melakukan klasifikasi sentimen positif dan negatif terhadap Tweet berbahasa Indonesia dengan mengimplementasikan algoritma *Deep Belief Network* dengan menghasilkan nilai keseluruhan akurasi 50.35%, *precision* 27%, *Recall* 50% dan *f1-Score* 35% [3]. Kemudian penelitian yang berjudul “Teknik Deep Learning untuk Analisis Profil Daerah Sehat di Indonesia” penelitian ini ditujukan memahami pentingnya Teknik deep learning dalam mengetahui opini masyarakat di media sosial[4]. Kemudian penelitian dengan judul “Studi Literatur tentang Perbandingan Metode untuk Proses Analisis Sentimen Di Twitter” penelitian ini membandingkan dua metode untuk masalah sentiment analisis yaitu Naïve Bayes dan SVM dengan hasil optimasi terbaik yaitu metode SVM[5].

Dari penelitian – penelitian sebelumnya yang sudah dijelaskan di atas masih perlu pengembangan, karena penelitian-penelitian sebelumnya hanya menggunakan satu algoritma dalam proses *analysis sentiment*. Perlu dilakukan penelitian yang dapat mengimplementasikan teknik *deep learning* dengan mengkombinasikan dua algoritma untuk proses analisis sentimen, kemudian menentukan hasil akurasi dari kombinasi dua algoritma tersebut. Topik penelitian yang akan diangkat ini berjudul **“KOMBINASI ALGORITMA CONVOLUTIONAL NEURAL NETWORK DAN LONG SHORT TERM MEMORY UNTUK MENINGKATKAN HASIL ANALISIS SENTIMEN”**.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan maka dapat dirumuskan permasalahannya yaitu bagaimana hasil akurasi dari kombinasi antara algoritma *Convolutional Neural Network* (CNN) dan algoritma *Long Short Time Memory* (LSTM) terhadap proses analisis sentimen berdasarkan data sosial media Twitter ?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini yaitu untuk mengkombinasikan algoritma *Convolutional Neural Network* (CNN) dengan algoritma *Long Short Time Memory* (LSTM) untuk meningkatkan hasil analisis sentimen berdasarkan data dari sosial media Twitter.

1.4 Manfaat Penelitian

1. Menerapkan metode *deep learning* dalam proses analisis sentimen.
2. Dapat digunakan untuk melakukan survey oleh pengampu kebijakan di perguruan tinggi yang ada di Indonesia.
3. Mengetahui akurasi prediksi dari kombinasi algoritma CNN dan LSTM.

1.5 Batasan Masalah

Berdasarkan permasalahan yang ada, agar masalah yang sedang ditinjau lebih terarah dan tercapai sasaran yang telah ditentukan, maka harus diberikan batasan-batasan masalah. Berikut merupakan batasan masalah dari proposal ini, yaitu:

1. Objek yang digunakan pada penelitian ini adalah data *tweet* dari pengguna Twitter di Kota Bandung.
2. *Tweet* yang digunakan yaitu *tweet* dengan berbahasa Indonesia saja.
3. Kata kunci untuk *Search* di Twitter menggunakan kata kunci *Informatika* dan *Teknik Informatika*.

4. Rentang waktu pengambilan data yang diambil yaitu dari tanggal 01 Januari 2015 sampai tanggal 30 Juli 2019.
5. Algoritma yang digunakan untuk penelitian adalah kombinasi dari algoritma *CNN (Convolutional Neural Network)* dan algoritma *LSTM (Long Short Time Memory)*.
6. Metode yang digunakan dalam penelitian adalah *deep learning*.
7. Persentase prediksi diambil secara keseluruhan sesuai dengan data yang ada.
8. *SDLC (System Development Life Cycle)* yang digunakan menggunakan model *Waterfall*.
9. Aplikasi yang dibuat menggunakan *command prompt* dan user interface sebagai media keluaran (*output*).

1.6 Sistematika Penulisan

Sistematika penulisan laporan tugas akhir ini disusun dalam beberapa bab yang masing-masing bab menguraikan beberapa pokok pembahasan. Adapun sistematika penulisan laporan ini yaitu sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisikan tentang latar belakang permasalahan yang diambil penulis, perumusan masalah yang dihadapi, batasan masalah, tujuan, state of the art, kerangka pemikiran, metodologi penelitian serta bagaimana sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini menjelaskan tentang konsep/teori apa saja yang berkaitan dengan topik yang diangkat, yang telah dibuat berdasarkan hasil penelitian dan hal-hal yang berguna dalam proses penulisan tugas akhir ini.

BAB III METODOLOGI PENELITIAN

Pada bagian ini merupakan bagian yang menjadi alur bagaimana penelitian dilakukan, mulai dari cara mengumpulkan data, mengolah atau menganalisis data dan menyimpulkan atau menetapkan simpulan dari sebuah hipotesis semua tercantum di dalam BAB 3.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan tentang pengujian sistem secara umum maupun spesifik. Pengujian sistem secara umum akan membahas mengenai lingkungan uji coba untuk menggunakan sistem ini. Selanjutnya secara lebih spesifik dijelaskan dalam pengujian sistem meliputi skenario pengujian beserta langkah - langkah dalam uji coba sistem untuk mengetahui aplikasi tersebut telah dapat menyelesaikan permasalahan yang dihadapi sesuai dengan yang diharapkan.

BAB V PENUTUP

Berisi tentang pernyataan singkat berupa kesimpulan dari pembahasan perangkat lunak yang dibuat secara keseluruhan dan saran untuk mengembangkan perangkat lunak yang lebih baik.