

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pertumbuhan pengguna Internet berbanding lurus terhadap perkembangan pengguna media sosial. Semakin besar pertumbuhan penggunaan Internet, semakin besar pula perkembangan pengguna media sosial. Pada tahun 2021, pengguna Internet khususnya media sosial di Indonesia mencapai kurang lebih 170 juta pengguna (Riyanto, 2021). Berbagai situs menyediakan fitur ulasan atau komentar dalam menilai layanan sebuah situs, aplikasi, film, dan lainnya. Hal tersebut menyebabkan jumlah data teks yang sangat banyak, entah melalui tulisan, komentar, maupun pesan mengenai sebuah topik.

Kumpulan data teks tersebut menjadi sumber data penting yang kaya pendapat, opini maupun sentimen. Analisis sentimen merupakan salah satu topik yang penting dalam kasus *Natural Language Processing* atau pemrosesan bahasa alami. Analisis sentimen adalah proses analisis pengelompokan bentuk-bentuk emosi, yang terdiri dari sentimen positif, negatif, maupun netral yang terdapat pada teks atau tulisan dengan penggunaan teknik analisis teks. Dilansir laman KBBI, analisis adalah penyelidikan terhadap suatu peristiwa untuk mengetahui keadaan sebenarnya. Sedangkan sentimen yaitu pendapat atau pandangan yang didasarkan pada perasaan yang berlebih-lebihan terhadap sesuatu. Analisis sentimen itu sendiri memiliki tujuan untuk menilai pendapat atau emosi terhadap sebuah produk, isu, maupun seorang tokoh masyarakat.

Salah satu algoritma yang dapat digunakan dalam menganalisis sentimen yaitu *Deep Learning*. *Deep learning* dapat memberikan performa yang lebih baik daripada *machine learning* dalam menganalisis data dalam jumlah yang besar (Rahmalia, 2021). *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM) merupakan algoritma *deep learning* yang cukup populer dan cocok digunakan dalam kasus analisis sentimen.

Kedua algoritma tersebut sudah banyak digunakan untuk mengerjakan berbagai masalah analisis sentimen.

Penelitian dilakukan oleh Yuliska, Qudsi, Lubis, Syaliman, & Najwa pada tahun 2021, mereka menganalisis sentimen pada data saran mahasiswa terhadap kinerja departemen di perguruan tinggi menggunakan CNN dengan *Word2Vec* sebagai *word embedding*-nya. *Simple CNN* dan *DoubleMax CNN* digunakan pada penelitian ini untuk mengetahui pengaruh kompleksitas arsitektur terhadap hasil klasifikasi sentimen. Hasil pengujian menunjukkan bahwa *DoubleMax CNN* dapat mengklasifikasi sentimen pada saran mahasiswa dengan sangat baik, yaitu mencapai akurasi tertinggi sebesar 98%, dengan nilai *precision* 98%, *F1-Score* 98%, dan *recall* sebesar 97%.

Sartini melakukan penelitian pada tahun 2020, yaitu analisis sentimen *twitter* bahasa Indonesia menggunakan algoritma CNN. Peneliti menggunakan 12 varian model CNN yang dikembangkan dengan kombinasi parameter yang berbeda pada *filter size*, *convolutional layers*, dan *number of filters* dan menggunakan *Word2Vec* sebagai *word embedding*-nya. Penelitian ini menggunakan pembagian data *training* dan data *testing* sebesar 80:20. Hasil yang didapat yaitu algoritma CNN mendapatkan akurasi yang lebih tinggi dibandingkan algoritma *machine learning* seperti KNN, SGD, dan SVM untuk analisis sentimen pada dataset berbahasa Indonesia. Algoritma CNN mendapat akurasi sebesar 81,4%, SGD 62,1%, KNN 52,3%, dan SVM 61,4%.

Nurrohmat & Azhari dalam penelitiannya menggunakan metode LSTM untuk analisis sentimen terhadap review novel pada tahun 2019. Tujuan penelitian ini adalah untuk melakukan klasifikasi terhadap review novel berbahasa Indonesia berdasarkan sentimen positif, negatif, dan netral dengan menggunakan metode LSTM. Data review novel berbahasa Indonesia yang diambil dari *website* goodreads.com merupakan dataset yang digunakan dalam penelitian ini. Metode LSTM juga akan dibandingkan dengan metode Naïve Bayes berdasarkan perhitungan dari nilai akurasi, *precision*, *recall*, dan *f-measure*. Dari hasil pengujian, diperlihatkan bahwa metode LSTM mendapatkan hasil akurasi yang lebih baik dibandingkan dengan metode

naïve bayes yaitu mendapatkan nilai akurasi 72.85%, *precision* 73%, *recall* 72%, dan *f-measure* 72%. Sedangkan hasil akurasi metode Naïve Bayes yaitu 67.88%, *recall* 68%, *precision* 69%, dan *f-measure* 68%.

Dari beberapa penelitian terdahulu, peneliti akan menganalisis perbandingan algoritma LSTM dengan CNN sehingga mendapatkan hasil kinerja yang lebih optimal dalam menganalisis sentimen. Masing-masing algoritma LSTM dan CNN akan menggunakan konfigurasi parameter yang terbaik saat dilakukan perbandingan dan dibantu dengan *open-source library* TensorFlow dan Keras.

1.2 Rumusan Masalah

Bagaimana perbandingan performa metode *Long Short-Term Memory* dengan *Convolutional Neural Network* dalam analisis sentimen?

1.3 Tujuan Penelitian

Membandingkan performa metode *Long Short-Term Memory* dengan *Convolutional Neural Network* dalam analisis sentimen sehingga dapat mengetahui metode mana yang lebih unggul dalam hal besarnya akurasi, kecepatan waktu latih, dan rendahnya nilai *loss*.

1.4 Batasan Masalah

1. Data bersumber dari situs Kaggle yang sudah dilabeli dengan label positif dan negatif. (Sumber: <https://www.kaggle.com/datasets/deniyulian/sentiment-analysis>)
2. Data merupakan *tweet* berbahasa Indonesia mengenai tayangan televisi di Indonesia yang ditujukan kepada acara Hitam Putih, Indonesia Lawyers Club, Kick Andy, dan Mata Najwa dengan jumlah sebanyak 400 data.
3. Metode yang akan dibandingkan adalah metode *Long Short-Term Memory* dan *Convolutional Neural Network*.

1.5 Luaran yang diharapkan

Luaran yang diharapkan dari penelitian ini yaitu untuk mendapatkan informasi bahwa metode mana yang performanya lebih baik dari metode *Long Short-Term Memory* dan *Convolutional Neural Network* dalam analisis sentimen.

1.6 Manfaat Penelitian

1. Menjadikan penelitian ini sebagai dasar untuk penelitian selanjutnya dalam menentukan metode terbaik untuk analisis sentimen.
2. Memperoleh informasi bahwa metode mana yang terbaik dari metode *Long Short-Term Memory* dan *Convolutional Neural Network* dalam analisis sentimen.

1.7 Sistematika Penulisan

BAB I : PENDAHULUAN

Bab 1 membahas latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, luaran yang diharapkan, batasan masalah, serta sistematika penulisan.

BAB II : TINJAUAN PUSTAKA

Bab 2 membahas uraian mengenai definisi dan teori analisis sentimen dan penggunaan metode secara umum.

BAB III : METODOLOGI PENELITIAN

Bab 3 membahas penjelasan metode yang dilakukan untuk menyelesaikan permasalahan yang menjadi objek penelitian

BAB IV : HASIL DAN PEMBAHASAN

Bab 4 membahas proses tahapan-tahapan penelitian, hasil penelitian, dan kendala-kendala yang ditemui selama proses penelitian.

BAB V : PENUTUP

Bab 5 berisi kesimpulan dan saran dari hasil penelitian yang dilakukan.

DAFTAR PUSTAKA

DAFTAR RIWAYAT HIDUP

LAMPIRAN