

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Instagram adalah salah satu media sosial yang paling banyak digunakan di Indonesia, dengan penggunaannya mencapai 79% dari rakyat Indonesia yang berusia 16-64 tahun [1]. Di balik kemajuan teknologi tersebut, tentunya cara orang berkomunikasi banyak mengalami perubahan, baik dalam mengkomunikasikan hal positif, maupun hal negatif, seperti *cyberbullying*. *Cyberbullying* merupakan sebuah perilaku berulang yang dilakukan untuk menakuti, membuat marah, atau mempermalukan korban lewat teks elektronik media sosial [2]. *Cyberbullying* dapat berdampak sangat meresahkan, bahkan orang yang mengalaminya dapat merasa dikucilkan, mengalami gangguan kesehatan fisik dan mental, depresi, dan yang paling parah, bunuh diri [3].

Media sosial memang sudah menyediakan layanan yang dapat digunakan untuk penanganan tindak *cyberbullying* dengan contohnya, melaporkan suatu akun, tetapi cara ini belum efektif karena penanganan suatu akun membutuhkan waktu lama. Masalah lain yang ditemukan adalah bahwa penanganan komentar berbau *cyberbullying* yang dilihat adalah, walaupun banyak macam bentuk *cyberbullying*, belum ada penanganan yang spesifik terhadap suatu bentuk *cyberbullying* tertentu.

Sebelumnya, telah dilakukan penelitian serupa tentang klasifikasi komentar *cyberbullying* dalam media sosial Twitter dengan menggunakan algoritma Naïve Bayes[4]. Algoritma Naïve Bayes dapat melakukan klasifikasi komentar *cyberbullying* dengan performa akurasi sebesar 87,67% dari 339.698 data. Dari penelitian lain yang telah diadakan pada tahun 2018 [5], *convolutional neural network* (CNN) juga diketahui memiliki performa yang baik dalam data dengan lokalitas yang besar, ketika kata-kata lebih banyak bobot tentang fitur-fitur di dalamnya, dan proyek tersebut sudah dapat dilihat sebagai langkah pertama dalam pembangunan *framework* untuk membangun aplikasi intervensi *cyberbullying*. Penelitian serupa juga telah dilakukan dengan komentar dalam media sosial Instagram, tapi dengan mendeteksi komentar spam yang telah dilakukan dengan metode *Complementary Naïve Bayes* dengan akurasi sebesar 92% dan dengan *Support Vector Machine* sebesar 87% juga telah dilakukan [6]. Penelitian lain klasifikasi komentar berbau *cyberbullying* dalam media

sosial Instagram dengan menggunakan *sentiment analysis* dan *Support Vector Machine* menunjukkan akurasi sebesar 90%[7]. Penelitian lain yang juga diadakan menggunakan *convolutional neural network* untuk mendeteksi kata-kata dengan indikasi *cyberbullying* memberikan akurasi 98.8%[8]. Berdasarkan latar belakang, diketahui bahwa diperlukan metode yang dapat melakukan klasifikasi ujaran atau komentar *cyberbullying* supaya penanganannya lebih mudah karena berdasarkan jenis-jenis *cyberbullying* tertentu. Berdasarkan penelitian-penelitian sebelumnya, diketahui bahwa *Convolutional Neural Network* dapat diterapkan untuk melakukan klasifikasi ujaran *cyberbullying*.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang sudah dibahas, ditentukan rumusan masalah sebagai berikut.

1. Bagaimana cara menerapkan *Convolutional Neural Network* untuk melakukan klasifikasi jenis *cyberbullying* dalam komentar Instagram?
2. Bagaimana kinerja dari implementasi *Convolutional Neural Network* dalam mengklasifikasikan jenis *cyberbullying* dalam komentar Instagram?

1.3 Batasan Permasalahan

Berdasarkan rumusan masalah, maka batasan masalah yang akan diteliti adalah sebagai berikut:

1. Penelitian ini akan melakukan deteksi pada lima jenis perilaku *cyberbullying*, yaitu berdasarkan *age* (usia), *ethnicity* (suku), *gender* (jenis kelamin), *religion* (agama), dan bukan *cyberbullying*[9].
2. Penelitian akan menggunakan dataset dari Kaggle.
3. Dataset yang akan digunakan adalah 47692 *tweets* dalam berbagai macam bahasa yang diambil dari Kaggle.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan dari diadakannya penelitian ini adalah sebagai berikut:

1. Menerapkan *convolutional neural network* untuk mengklasifikasikan jenis *cyberbullying* dalam komentar Instagram.
2. Mengukur kinerja dari implementasi *convolutional neural network* untuk mengklasifikasikan komentar jenis *cyberbullying* dalam komentar Instagram.

1.5 Manfaat Penelitian

Manfaat yang didapat melalui penelitian ini adalah mengembangkan suatu metode yang dapat melakukan klasifikasi komentar Instagram yang memiliki indikasi *cyberbullying* dengan akurat. Manfaatnya bagi masyarakat adalah, dapat membantu masyarakat lebih tahu jenis-jenis *cyberbullying*

1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
Berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.
- Bab 2 LANDASAN TEORI
Berisi landasan-landasan teori terkait topik penelitian.
- Bab 3 METODOLOGI PENELITIAN
Berisi metodologi penelitian yang dilakukan serta perancangan sistem berupa *flowchart*.
- Bab 4 HASIL DAN DISKUSI
Berisi spesifikasi sistem, implementasi sistem, pengujian sistem, dan hasil evaluasi.
- Bab 5 SIMPULAN DAN SARAN
Berisi simpulan dari penelitian yang telah dilakukan dan saran untuk penelitian selanjutnya.