

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Media sosial merupakan sebuah media *online*, dimana para penggunanya bisa dengan mudah berpartisipasi, berbagi, dan menciptakan isi meliputi blog, jejaring sosial, wiki, forum dan dunia virtual (Cahyono, 2016). Media sosial bisa jadi tempat berinteraksi yang positif ataupun negatif. Salah satu contoh negatif dari media sosial adalah *cyberbullying*. *Cyberbullying* atau perundungan secara daring merupakan sebuah kegiatan dengan sengaja untuk mengirimkan pesan teks elektronik kepada pengguna lain dengan tujuan untuk mengejek, melecehkan, mengancam, dan mengganggu (Mulya, 2017).

Menurut Ditch The Label yang merupakan sebuah organisasi lembaga donasi *anti-bullying*, Instagram merupakan media sosial yang sering digunakan untuk melakukan *cyberbullying*. Hal itu didasarkan pada survei kepada 10.020 remaja asal Inggris dengan kisaran usia 12 hingga 20 tahun, dari hasil survei tersebut 42 persen responden mengaku pernah menjadi korban *cyberbullying* di Instagram (Pratama, 2021). Polda Metro Jaya menuturkan bahwa paling tidak terdapat 25 permasalahan *cyberbullying* yang dilaporkan setiap harinya. Pada tahun 2018 Komisi Perlindungan Anak Indonesia mengatakan jumlah anak korban *bullying* mencapai 22,4%. Tingginya nilai itu dipicu oleh banyaknya konsumsi internet pada anak (Putra, 2019). Oleh karena itu, permasalahan *cyberbullying* menarik untuk dikaji sebagai topik penelitian.

Diperlukan sebuah sistem yang dapat mengklasifikasikan komentar yang bersifat negatif (*cyberbullying*) atau positif (mendukung). Berdasarkan penelitian Eric Darson yang dilakukan pada tahun 2019 menyimpulkan bahwa dalam mengklasifikasikan komentar *cyberbullying* menggunakan Support Vector Machine mendapatkan nilai akurasi 76.77% (Darson, 2019). Dalam penelitiannya juga, Darson (2019) menyarankan untuk penelitian selanjutnya klasifikasi komentar *cyberbullying* dapat menggunakan metode klasifikasi lainnya. Salah satu metode klasifikasi lainnya adalah metode XGBoost. Penelitian yang dilakukan oleh Babacara Gaye dan Aziguli Wulamu menyatakan bahwa algoritma XGBoost memiliki akurasi dan performa yang lebih tinggi dibandingkan dengan algoritma Support Vector Machine (Gaye & Wulamu, 2019). XGBoost dapat memiliki performa lebih baik dari Support Vector Machine dikarenakan XGBoost merupakan metode *tree ensemble* yang telah dioptimasi baik secara sistem maupun algoritmanya (Morde, 2019).

Terdapat penelitian yang serupa mengenai implementasi XGBoost pada sentimen analisis dalam media sosial Facebook yang dilakukan oleh Ega Pamungkas pada tahun 2020 yang menghasilkan akurasi 74.8%, *precision* 50%, *recall* 48%, dan *F1 Score* sebesar 49% (Pamungkas, 2020). Selain itu, penelitian yang dilakukan oleh Saifullah dkk. (2021) menyatakan bahwa XGBoost yang menggunakan ekstraksi fitur TF-IDF memiliki performa akurasi dan presisi yang sedikit lebih baik dari ekstraksi fitur Count-Vectorization.

Berdasarkan penelitian-penelitian yang telah disebutkan, XGBoost merupakan salah satu metode yang cocok untuk mengklasifikasikan data. Oleh karena itu

XGBoost dipilih sebagai metode untuk mengembangkan sistem klasifikasi komentar *cyberbullying*.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan sebelumnya, maka rumusan masalah yang diangkat pada penelitian ini adalah sebagai berikut.

1. Bagaimana menerapkan algoritma Extreme Gradient Boosting untuk klasifikasi komentar *cyberbullying* ?
2. Bagaimana mengukur performa klasifikasi komentar *cyberbullying* yang menggunakan algoritma Extreme Gradient Boosting?

1.3 Batasan Masalah

Batasan masalah dari penelitian ini akan diurai dalam poin-poin sebagai berikut.

1. Lingkup bahasa adalah bahasa Indonesia sesuai dengan KBBI.
2. Terdapat dua kelas klasifikasi yaitu *bullying* dan *non-bullying*.
3. *Dataset* yang digunakan diambil dari penelitian Luqyana (2018) berjudul Komentar Instagram Cyberbullying dan Andriansyah (2017) berjudul Komentar Selebgram Cyberbullying.
4. Jika terdapat *emoticon* pada komentar maka *emoticon* akan dihapus.

1.4 Tujuan Penelitian

Tujuan penelitian akan dijabarkan dalam poin-poin sebagai berikut.

1. Mengimplementasikan algoritma Extreme Gradient Boosting untuk klasifikasi komentar *cyberbullying*.

2. Mengukur performa algoritma Extreme Gradient Boosting untuk klasifikasi komentar *cyberbullying*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini akan dijabarkan dalam poin-poin sebagai berikut.

1. Memberikan fasilitas kepada pengguna untuk mengetahui apakah komentar yang dimasukkan bermuatan *cyberbullying* atau tidak.
2. Memberikan fasilitas kepada peneliti yang memiliki penelitian sejenis dalam memberikan hasil prediksi berupa label pada *dataset*.

1.6 Sistematika Penulisan

Sistematika penulisan pada skripsi ini terbagi menjadi 5 bagian yang akan dijabarkan sebagai berikut.

1. BAB 1 PENDAHULUAN

Pada Bab ini membahas latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, batasan penelitian, dan sistematika penulisan.

2. BAB 2 LANDASAN TEORI

Pada Bab ini menjabarkan tentang teori-teori yang digunakan dalam penelitian ini.

3. BAB 3 METODOLOGI PENELITIAN

Pada Bab ini berisi langkah-langkah serta perancangan yang dilakukan pada penelitian ini.

4. BAB 4 HASIL DAN DISKUSI

Pada Bab ini berisi hasil dari penelitian yang telah dilakukan

5. BAB 5 KESIMPULAN DAN SARAN

Pada Bab ini berisi mengenai jawaban atas tujuan penelitian yang telah dikemukakan pada Bab 1. Bab ini juga berisi saran untuk penelitian lebih lanjut.

UMMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA