

BAB I PENDAHULUAN

1.1 Latar Belakang

Pesatnya perkembangan website membuatnya tercipta banyak sosial media seperti Facebook, Instagram, Tiktok, dan Twitter. Twitter di Indonesia menjadi salah satu media sosial yang sering digunakan. hal tersebut dibuktikan dengan data dari *We Are Social* yang menyatakan pengguna Twitter Indonesia mencapai 14,05 juta pada tahun 2021, sedangkan menurut Kementerian Komunikasi dan Informatika pengguna Twitter di Indonesia menempati urutan kelima dunia. Pengguna Twitter dapat berbagi informasi hingga menyampaikan pendapatnya terkait hal – hal tertentu tak jarang pengguna juga berbagi hal – hal yang bersifat pribadi melalui *tweets*.

Hadirnya media sosial bersamaan dengan dampak positif dan negatif yang ditimbulkannya. Dampak positif dari media sosial adalah kemudahan dalam berbagi serta mendapatkan informasi seperti pada media sosial pada Twitter, sedangkan dampak negatif yang ditimbulkan media sosial adalah media sosial digunakan sebagai sarana untuk melakukan *bullying* [1]. *Bullying* atau perundungan adalah kegiatan yang dilakukan secara konstan serta berkali – kali dan merugikan fisik maupun psikis oleh orang lain [2], sedangkan *cyberbullying* adalah tindakan penyerangan berdasarkan tingkah laku *bullying* yang dilakukan oleh orang lain pada media sosial [3].

Di Indonesia menurut UNICEF (*United Nations Children's Fund*) melalui U-Report, melakukan survei dengan responden sebanyak 2777 dan ditemukan bahwa 45% dari mereka pernah menjadi korban *cyberbullying* serta sebanyak 71% *cyberbullying* terjadi pada sosial media, seperti Twitter. Pada Twitter aktivitas *bullying* atau perundungan dilakukankan secara cepat serta sangat agresif menyerang korban dengan cacian yang tidak terdeteksi oleh sistem Twitter [4]. Di sisi lain dampak dari tindakan *bullying* baik secara *cyber* atau tradisional dikaum remaja mampu berakibat depresi, tindakan bunuh diri dan percobaan bunuh diri [5].

Cyberbullying memiliki dampak yang berbahaya, oleh sebab itu maka dibutuhkan upaya deteksi atau pencegahan supaya tidak terjadi bahaya pada korbannya [5]. Dalam melakukan deteksi aktivitas *cyberbullying* pada twitter maka dibutuhkan *machine learning* dengan menggunakan algoritma *K-Nearest Neighbor* (KNN) yang mampu mendeteksi atau membedakan *tweets bullying* dengan *tweets* bukan *bullying*. *K-Nearest Neighbor* merupakan metode klasifikasi objek berdasarkan jarak terdekat dari data pembelajaran. Dalam mengklasifikasikan dataset yang digunakan berasal dari *tweets* pada media sosial *Twitter*, dataset tersebut akan dibagi kedalam dua kelas atau *class*, yaitu *bullying* dan bukan *bullying*. Kelas *bullying* ditujukan untuk *tweets* yang isinya tindakan *bullying*, sedangkan kelas bukan *bullying* ditujukan untuk *tweets* yang isinya tidak ada tindakan *bullying*.

1.2 Rumusan Masalah

Berdasarkan latarbelakang yang telah dikemukakan sebelumnya maka dapat dirumuskan beberapa rumusan masalah berikut :

1. Bagaimana kinerja algoritma *K-Nearest Neighbor* diterapkan dalam klasifikasi teks ?
2. Bagaimana penerapan algoritma *K-Nearest Neighbor* dalam klasifikasi *tweets bullying* maupun bukan *bullying* ?

1.3 Batasan Masalah

Pada penelitian ini juga memiliki batasan masalah agar penelitian bisa terfokuskan pada tujuan penelitian. Berikut beberapa batasan masalah yang terdapat pada penelitian ini :

1. Metode yang digunakan untuk mengolah data adalah menggunakan algoritma *K-Nearest Neighbor* (KNN).
2. Dataset yang digunakan merupakan data yang berasal dari Twitter dan berbentuk teks.
3. Proses pelabelan data dilakukan secara manual berdasarkan pemikiran pribadi tanpa adanya bantuan dari ahli.

1.4 Tujuan Penelitian

Penelitian ini memiliki tujuan untuk membuat serta menghasilkan permodelan yang mampu mendeteksi *tweets* yang memiliki unsur *bullying* serta tidak memiliki unsur *bullying* dengan menggunakan algoritma *K-Nearest Neighbor*. Tujuan lain dari penelitian adalah untuk menguji kinerja dari *K-Nearest Neighbor* dalam mengolah data yang berasal dari Twitter.

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan mampu dimanfaatkan sebagai acuan dalam pengembangan sistem klasifikasi teks yang menggunakan metode *K-Nearest Neighbor*. Selain itu diharapkan hasil penelitian ini mampu memberikan kontribusi dalam pengembangan metode yang digunakan yaitu metode *K-Nearest Neighbor* serta mampu memberikan kontribusi terhadap asal dari data yang digunakan yaitu Twitter.

1.6 Sistematika Penulisan

Berikut ini adalah merupakan sistematika penulisan pada skripsi ini :

1. BAB I PENDAHULUAN, berisi Latar Belakang, Rumusan Masalah, Tujuan Masalah, Manfaat Penelitian serta Sistematika Penulisan. Secara garis besar bab I ini menjelaskan secara umum mengenai penelitian yang akan dilakukan peneliti.
2. BAB II TINJAUAN PUSTAKA, bagian ini terdapat Studi Literatur dan Dasar Teori. Bagian ini menjelaskan tentang jurnal, skripsi, dan paper yang dijadikan rujukan pada skripsi ini, serta berisi mengenai dasar teori yang digunakan pada skripsi.
3. BAB III METODE PENELITIAN, berisi Objek Penelitian, Alur Penelitian serta Alat dan Bahan, bagian ini menjelaskan tentang data yang digunakan dan langkah – langkah penelitian.
4. BAB IV HASIL DAN PEMBAHASAN, bagian ini secara garis besar menjelaskan hasil penelitian, yang telah dilakukan.

5. BAB V PENUTUP, terdapat dua hal yaitu, Kesimpulan yang berisi simpulan dari penelitian yang telah dilakukan, serta Saran yang menjelaskan masalah – masalah atau kekurangan yang ditemukan selama penelitian.

