

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Di era yang semakin maju, media sosial juga semakin berkembang. Salah satu media sosial yang populer dan berkembang pesat adalah Instagram. Menurut survei yang dilakukan oleh *Ditch The Label* pada tahun 2017 kepada 10.020 siswa berusia antara 12 sampai 20 tahun pengguna aktif Instagram mencapai 78% atau sama dengan 7816 pengguna dari total siswa (*Ditch The Label*, 2017), sayangnya masih banyak pengguna yang tidak bijak dalam menggunakan media Instagram. Banyak pengguna yang menyalahgunakan dan tidak beretika dalam bersosialisasi melalui Instagram.

Permasalahan tersebut menyebabkan timbulnya tindakan *cyberbullying*. *Cyberbullying* merupakan bentuk penindasan yang dilakukan oleh seseorang yang bertujuan untuk melecehkan korbannya melalui media sosial (Terry Brequet, 2010). *Cyberbullying* dapat berupa pesan atau gambar yang kemudian disebar dengan tujuan mempermalukan korban bagi pengguna lain yang melihatnya (Rahayu, 2012).

Menurut survei yang dilakukan oleh lembaga *anti-bullying*, *Ditch The Label* dapat diketahui bahwa media sosial Instagram menjadi media sosial yang paling umum digunakan dalam melakukan *cyberbullying*. Hasil menunjukkan lebih dari 42% korban *cyberbullying* mendapatkannya dari media Instagram (*Ditch the label*, 2017). *Cyberbullying* menjadi masalah yang serius yang dapat memberikan dampak buruk terhadap psikologis korban bahkan sampai melakukan bunuh diri, sehingga permasalahan tersebut harus segera ditangani terutama pada remaja.

Segala bentuk *bullying* yang dilakukan di media sosial merupakan salah satu bentuk tindak pidana penyalahgunaan sosial media yang diatur dalam Undang – Undang Informasi dan Transaksi Elektronik (UU ITE) nomor 11 tahun 2008 pasal 27 ayat 3 yang berbunyi “*Setiap Orang dengan sengaja dan tanpa hak mendistribusikan dan/atau mentransmisikan dan/atau membuat dapat diaksesnya Informasi Elektronik dan/atau Dokumen Elektronik yang memiliki muatan penghinaan dan/atau pencemaran nama baik*”.

Permasalahan *Cyberbullying* pada komentar Instagram menjadi hal yang penting dan menarik untuk dikaji sebagai pemrosesan data teks. Salah satu jenis pemrosesan data teks adalah analisis sentimen. Analisis sentimen merupakan sebuah proses yang digunakan untuk mengolah komentar pengguna dari media Instagram. Analisis sentimen dibagi menjadi 3 jenis opini yaitu opini negatif, positif dan netral. Dalam melakukan analisis sentimen dibutuhkan metode untuk menunjang klasifikasi.

*SVM*, *Naïve Bayes*, dan *Random Forest* merupakan algoritma *machine learning* yang paling umum digunakan dalam klasifikasi teks, maka dari itu penulis ingin membandingkan 3 algoritma tersebut. Berdasarkan penelitian sebelumnya yang dilakukan oleh Maria Ismiati pada tahun 2018 telah dilakukan deteksi komentar negatif di Instagram dengan menggunakan algoritma *Naïve Bayes* dengan objek  $\pm 50$  komentar Instagram yang diambil secara acak. Dari penelitian tersebut menghasilkan nilai akurasi sebesar 76,7% (Maria, 2018). Penelitian lainnya dilakukan oleh Luqyana, Cholissodin dan Perdana pada tahun 2018 telah melakukan analisis sentimen *Cyberbullying* komentar Instagram dengan menggunakan metode *SVM* yang kemudian menghasilkan nilai akurasi sebesar 90% (Luqyana, et al, 2018). Penelitian lainnya dilakukan oleh Asep, Warih dan Anisa yang telah melakukan Analisis sentimen dan peringkasan pada ulasan produk menggunakan algoritma *Random Forest* yang kemudian menghasilkan akurasi sebesar 75% (Asep, Warih & Anisa).

Alasan penulis membandingkan 3 algoritma *Naïve bayes*, *SVM* dan *Random Forest* karena banyak penelitian sebelumnya yang membahas tentang klasifikasi teks dengan menggunakan algoritma *SVM*, *Naïve Bayes* dan *Random Forest* dengan hasil akurasi yang baik. Pada penelitian ini juga menggunakan parameter tuning untuk menunjang performa tiap algoritma. Berdasarkan uraian yang sudah dijelaskan, diharapkan dapat mengetahui algoritma yang paling cocok untuk diterapkan dalam klasifikasi sentimen pada komentar Instagram dan dapat mengklasifikasi teks berdasarkan kelas negatif, positif dan netral.



## 1.2 Perumusan Masalah

Berdasarkan latar belakang diatas, permasalahan yang akan dikaji dalam penelitian ini adalah sebagai berikut :

- a. Bagaimana mengklasifikasi teks berdasarkan kelasnya (negatif, positif dan netral) ?
- b. Bagaimana mengetahui algoritma yang paling cocok untuk klasifikasi sentimen pada media Instagram?

## 1.3 Batasan Masalah

Agar penelitian yang dilakukan tidak menyimpang dari pokok pembahasan dan mencapai sasaran, maka diperlukan batasan sebagai berikut :

- a. Data diambil dalam rentang waktu postingan bulan Maret – Juni 2020.
- b. Data yang digunakan dalam penelitian ini adalah data komentar pada beberapa akun Instagram.
- c. Data teks diklasifikasikan ke dalam tiga kelas yaitu negatif, positif dan netral.
- d. Data yang digunakan hanya data komentar Instagram yang berupa teks, tidak mengandung gambar.
- e. Bahasa pemrograman yang digunakan dalam membangun sistem adalah bahasa pemrograman *Python*.

## 1.4 Manfaat Penelitian

Manfaat yang diharapkan dalam penelitian ini adalah sebagai berikut :

- a. Diharapkan dapat menjadi referensi bagi peneliti selanjutnya, khususnya yang berkaitan dengan pengklasifikasian teks.
- b. Diharapkan dapat bermanfaat bagi pengembangan ilmu komunikasi dalam hal pengembangan media sosial agar tidak terjadi tindakan *cyberbullying*.
- c. Diharapkan dapat bermanfaat bagi orang-orang khususnya remaja dan dewasa agar lebih bijak dalam menggunakan sosial media dan tidak melakukan tindakan *cyberbullying* pada media sosial.

- d. Diharapkan dapat membantu dalam mengklasifikasi teks berdasarkan kelasnya.
- e. Dapat membantu mengetahui algoritma yang paling cocok untuk analisis sentimen pada komen Instagram.

### **1.5 Tujuan Penelitian**

Berdasarkan latar belakang diatas, tujuan dari penelitian ini adalah sebagai berikut :

- a. Mengklasifikasi sentimen pada komentar Instagram berdasarkan kelas negatif, positif dan netral.
- b. Mengetahui algoritma yang paling cocok untuk analisis sentiman pada komentar Instagram.

