

## **BAB 1. PENDAHULUAN**

### **1.1 Latar Belakang**

Internet adalah suatu jaringan komunikasi yang memiliki fungsi untuk menghubungkan antara satu media elektronik dengan media elektronik lainnya dengan cepat dan tepat. Internet telah berkembang sangat pesat dari waktu pertama kali terbentuknya internet. Dari tahun ke tahun perkembangan internet mengalami perubahan yang signifikan, dari segi cakupan, banyak negara berlomba-lomba memperluas jaringannya dengan menggunakan satelit. Dari sisi kecepatan pun, banyak industri teknologi yang mulai mengembangkan generasi jaringan. Dimulai dari 2G, 3G, 4G, hingga yang paling baru saat ini adalah 5G. Penggunaan internetpun memiliki perkembangan yang besar. Jika pada awal penggunaannya internet hanya digunakan untuk kebutuhan riset, pada era digital seperti saat ini penggunaan internet telah mencapai ranah pendidikan, sosial budaya, politik, bisnis dan lain sebagainya.

Penggunaan internet di Indonesia mengalami peningkatan yang sangat pesat. Berdasarkan data dari layanan manajemen konten HootSuite, pengguna internet di Indonesia pada awal 2021 mencapai 202,6 juta jiwa. Jumlah pengguna internet ini meningkat 15,5% atau 27 juta jiwa jika dibandingkan pada Januari 2020 lalu. Total jumlah penduduk Indonesia saat ini adalah 274,9 juta jiwa, ini berarti, penetrasi internet di Indonesia pada awal 2021 mencapai 73,7% (We Are Social & HootSuite, 2021). HootSuite juga melaporkan bahwa pengguna internet berusia 16 hingga 64 tahun diketahui memiliki beberapa perangkat elektronik berbeda, antara lain telepon genggam, laptop/PC, tablet, dan sebagainya. Berdasarkan survei yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia, mayoritas pengguna mengakses internet selama lebih dari 8 jam dalam satu hari terutama sewaktu pandemi COVID-19. Terdapat lima alasan untuk masyarakat mengakses internet di Indonesia, yaitu pengaksesan media sosial, pesan dan komunikasi, bermain *game online*, dan pembelian barang secara *online* atau belanja *online*.

Mayoritas masyarakat Indonesia banyak mengakses layanan internet dengan menggunakan smartphone. Platform media sosial yang paling banyak digunakan oleh masyarakat Indonesia adalah Facebook, Instagram dan Twitter.

Instagram adalah sebuah platform media sosial yang saat ini berada di bawah naungan perusahaan yang dimiliki oleh Mark Zuckerberg, yaitu Facebook. Berdasarkan data dari statista, penggunaan Instagram termasuk kedalam lima besar media sosial paling populer di dunia terhitung per bulan april tahun 2021 dengan 1,287 miliar pengguna dan terus meningkat setiap tahunnya. Pengguna Instagram berdasarkan demografi gender, wanita adalah pengguna terbanyak mencapai 51,1% di bulan januari dan laki-laki adalah 48,9% pengguna di bulan april. Berdasarkan demografi usia, pengguna terbanyak berada pada rentang usia 18 – 34 tahun mencapai 64% dari total pengguna dan hanya 30% saja pengguna yang memiliki rentang usia diatas 35 tahun. Indonesia termasuk kedalam 5 besar dengan pengguna Instagram terbanyak di dunia, mencapai 86 juta pengguna. Dari 86 juta pengguna tersebut, yang terbanyak berada pada rentang usia 18 – 34 tahun dengan 64,8% dari total pengguna Instagram di Indonesia.

Banyaknya pemuda yang menggunakan media sosial saat ini, tentunya menimbulkan suatu fenomena baik itu berupa dampak positif maupun dampak negatif. Contoh dari dampak positif penggunaan media sosial misalnya adalah dapat berinteraksi dengan teman dan keluarga yang jauh, mengasah keterampilan baru yang terlihat di media sosial, dapat dengan mudah membangun relasi dengan orang yang tidak kita kenal sebelumnya, dan bisa mendapatkan penghasilan dari keterampilan yang diunggah di media sosial. Dampak negatif dari penggunaan media sosial contohnya antara lain meningkatnya gangguan kecemasan, apatis terhadap lingkungan, dan maraknya *cyberbullying*.

Fenomena maraknya *cyberbullying* karena bebasnya dalam berkomentar dan tidak adanya norma dalam bermedia sosial. Instagram menjadi media sosial yang memiliki tingkat *cyberbullying* tertinggi di dunia dengan persentase mencapai 42%, kemudian Facebook menjadi media sosial

dengan *cyberbullying* tertinggi kedua dan Snapchat peringkat ketiga dengan masing-masing persentase sebesar 37% dan 31% dari total 10.020 remaja asal Inggris yang disurvei (Sheri Bauman, dkk). Tentunya *cyberbullying* memiliki banyak efek negatif kepada orang yang mengalami perundungan, antara lain merasakan kegelisahan, depresi, menyakiti diri sendiri, bahkan sampai melakukan bunuh diri. Salah satu hal yang dapat dikategorikan sebagai *cyberbullying* adalah melakukan sarkasme atau sindiran.

Menurut Tarigan (2013), sarkasme berasal dari Bahasa Yunani yaitu *sarkasmos* yang berasal dari kata kerja *sakasein* yang berarti merobek robek seperti anjing, menggigit bibir karena marah atau bicara dengan kepahitan. Gaya Bahasa sarkasme termasuk kedalam Bahasa pertentangan, sehingga dapat diartikan sebagai pemilihan kata yang memiliki arti yang bertentangan (Anshari & Hafiz, 2018).

Penggunaan kalimat sarkas atau sindiran saat berkomentar di sosial media, terutama Instagram dapat mengakibatkan salah paham karena tidak adanya interaksi antara penulis komentar dengan orang yang dikomentari. Bisa jadi dalam sebuah komentar tersebut hanya ingin melakukan kritik dan penulis komentar tidak berniat melakukan perundungan. Para pengguna media sosial biasanya melakukan penggunaan kalimat sarkasme untuk mengkritik sebuah fenomena yang baru atau sedang terjadi, tetapi ungkapan yang digunakan sering menyeleweng dari aturan kesantunan berbahasa. Instagram adalah salah satu media sosial yang memiliki ruang yang tak terbatas, sehingga semua orang bebas menggunakannya. Kebebasan yang dimiliki Instagram membuat orang-orang memiliki keberanian untuk mengungkapkan perasaan mereka dan banyak dari mereka yang menggunakan ungkapan sarkasme sehingga seringkali menimbulkan pelanggaran kesantunan dalam berbahasa.

Banyaknya pelanggaran dalam kesantunan berbahasa pada media sosial Instagram menjadikan penulis menyadari bahwa dibutuhkan sebuah sistem agar bisa mendeteksi kalimat sarkasme dalam sebuah *post* yang diunggah dalam media sosial ini. Pendeteksian kalimat sarkasme dapat membuat kesantunan dalam berbahasa saat berkomentar menjadi lebih tinggi.

*Sentimen analysis* atau biasa disebut dengan *Opinion mining* adalah riset komputasional dari opini, sentimen, dan emosi yang dituangkan secara tekstual lalu diklasifikasikan menjadi kelompok sentimen positif dan negatif (Feizar, Indrianti, & Yudistira, 2014). Komentar pada Instagram pun dapat diklasifikasikan berdasarkan sentimennya. Dengan menggunakan analisis sentimen, data dapat dikelompokkan menjadi sentimen positif dan negatif. Tetapi analisis sentimen kurang akurat dalam mendeteksi adanya sarkasme, karena pada dasarnya sarkasme adalah kalimat yang mengekspresikan perasaan negatif dengan kata positif yang mengakibatkan perubahan polaritas kalimat dari positif ke negatif maupun sebaliknya.

Dalam melakukan analisis sentimen maka dibutuhkan metode untuk menunjang aktifitas klasifikasi. Berdasarkan penelitian yang dilakukan oleh Wanda, Imam Cholissodin dan Rizal Setya Perdana (2018) untuk analisis sentimen komentar pada Instagram dengan mengimplementasikan metode klasifikasi Support Vector Machine (SVM) menghasilkan akurasi tertinggi sebesar 90%. Kemudian berdasarkan penelitian yang dilakukan Debby Alita, Sigit Priyanta dan Nur Rokhman (2019), untuk deteksi sarkasme pada twitter menggunakan dua tahap. Tahap pertama untuk analisis sentimen dan tahap kedua untuk deteksi sarkasme. Pada tahap analisis sentimen peneliti mengimplementasikan metode SVM dan Naïve Bayes dengan masing-masing akurasi tertinggi adalah 77,79% dan 51,56% yang menunjukkan bahwa SVM memiliki akurasi yang lebih tinggi dibandingkan Naïve Bayes. Kemudian untuk deteksi sarkasme penelitian tersebut menerapkan tiga metode yaitu Naïve Bayes, SVM, dan Random Forest Classifier. Hasil penelitian tersebut mendapatkan nilai akurasi tertinggi masing-masing adalah 54.88%, 60.34% dan 60.61%.

Berdasarkan data-data yang telah disebutkan, maka penelitian ini akan menggunakan dua tahap. Tahap pertama untuk analisis sentimen akan menerapkan TF-IDF (*Term Frequency-Inverse Document Frequency*) untuk pembobotan kata dan SVM untuk analisis sentimennya. Tahap kedua adalah

menganalisis akurasi dari metode Logistic Regression untuk deteksi sarkasme pada komentar Instagram.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang diatas, maka dapat dirumuskan masalah dalam penelitian ini yaitu :

1. Bagaimana melakukan klasifikasi terhadap komentar instagram sehingga mengarah pada *bullying*?
2. Bagaimana mendeteksi sarkasme pada teks yang telah diklasifikasikan sebagai *bullying*?
3. Apakah *Logistic Regression* dapat digunakan dalam melakukan deteksi kalimat sarkasme?

## **1.3 Tujuan**

Dalam penelitian ini, tujuan yang ingin dicapai oleh penulis adalah melihat jika metode *Logistic Regression* dapat digunakan untuk deteksi sarkasme terhadap komentar pada *post* di Instagram.

## **1.4 Manfaat**

Dengan adanya penelitian ini diharapkan dapat menghasilkan sistem yang dapat mendeteksi sarkasme pada teks yang terindikasi *Bullying* pada komentar di Instagram.