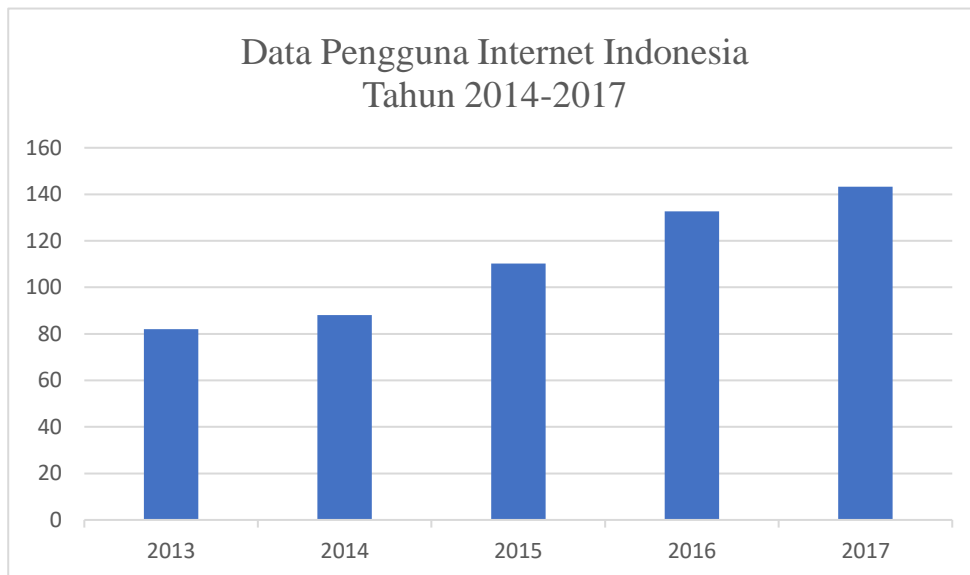


BAB I

PENDAHULUAN

1.1 Latar Belakang

Dewasa ini, perkembangan teknologi internet semakin meningkat dengan pesat. Dengan adanya internet, memungkinkan pengguna untuk selalu terhubung dan berinteraksi dengan pengguna internet lainnya secara langsung tanpa dibatasi ruang dan waktu. Hingga saat ini teknologi internet telah diterapkan oleh berbagai negara di belahan dunia manapun, salah satunya adalah Indonesia. Bahkan Indonesia termasuk negara yang paling cepat pertumbuhan pengguna internet sebagaimana survei yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) pada tahun 2017 [1]. Dari survei tersebut didapatkan hasil bahwa terjadi peningkatan pengguna internet Indonesia yang signifikan tiap tahunnya seperti yang dapat dilihat pada Tabel 1.1. Dan untuk tahun 2019 ini diproyeksikan menembus angka 175 juta, atau sekitar 65,3% dari total penduduk 268 juta [2].



Gambar 1.1 Data pengguna internet Indonesia dalam satuan juta

Namun bagai pisau bermata dua, selain memberikan dampak positif internet juga memberikan dampak negatif bila terjadi penyalahgunaan layanan tersebut. Salah

satu dampak negatif akibat penyalahgunaan internet adalah *cyberbullying*. *Cyberbullying* merupakan bentuk intimidasi, pelecehan atau perlakuan kasar secara verbal ditujukan kepada korban secara terus menerus yang dilakukan di dunia maya [3]. Berdasarkan hasil survei UNICEF Indonesia pada tahun 2016, 41-51% remaja yang berusia 13-15 tahun pernah menjadi korban *cyberbullying* [4]. Dampak yang ditimbulkan oleh tindakan *cyberbullying* ini lebih menyakitkan daripada *bullying* yang dilakukan secara fisik. Korban *cyberbullying* akan mengalami tekanan berat, depresi, bahkan melakukan tindakan yang lebih fatal yaitu bunuh diri. Tindakan *cyberbullying* umumnya dilakukan melalui media sosial. Internet dan media sosial mengakibatkan jumlah kasus *cyberbullying* meningkat karena karakteristiknya yang memungkinkan orang menyebarkan informasi dengan mudah dan cepat [5]. Salah satu media sosial yang paling banyak digunakan adalah Instagram.

Instagram merupakan salah satu media sosial terpopuler di dunia yang memungkinkan setiap pengguna untuk berbagi kiriman foto, gambar dan video [6]. Instagram saat ini telah memiliki banyak pengguna aktif dari seluruh kalangan mulai dari pengguna umum, artis, publik figur hingga petinggi negara. Indonesia menempati urutan ketiga negara dengan pengguna aktif Instagram sebesar 55 juta [7]. Setiap pengguna Instagram juga dapat memberikan *like* dan komentar terhadap kiriman yang telah dibagikan. Namun bila dikaitkan dengan fenomena *cyberbullying*, Instagram masih belum menyediakan fitur untuk mendeteksi komentar *cyberbullying*. Fitur yang disediakan oleh Instagram adalah sebatas meminimalisir tindakan *cyberbullying* seperti pembatasan jumlah komentar atau menonaktifkan kolom komentar pada suatu postingan, fitur "*hide inappropriate comments*" yang berguna untuk menyembunyikan komentar berdasarkan kata kunci yang disediakan oleh Instagram, dan fitur *report* yang digunakan untuk melaporkan suatu komentar ke pihak Instagram [8]. Pengguna Instagram juga dapat mengunci akunnya menjadi *private* untuk menjaga privasi setiap postingan. Salah satu teknik untuk mengolah data yang berupa teks adalah *text mining*.

Text mining bertujuan untuk mencari informasi yang sebenarnya secara eksplisit dapat dilihat dan dimengerti oleh manusia menggunakan algoritma yang diolah oleh komputer tanpa menghilangkan data penting yang ada [9]. Salah satu

penerapan *text mining* di dunia nyata adalah klasifikasi teks. Pada penelitian ini, klasifikasi teks dilakukan untuk mengklasifikasikan suatu komentar pengguna Instagram ke dalam kelas *cyberbullying* dan bukan *cyberbullying*. Oleh karena itu, diperlukan suatu algoritma yang mampu melakukan klasifikasi ke dalam dua kelas tersebut. Terdapat beberapa algoritma yang dapat digunakan untuk pengklasifikasian seperti *Support Vector Machine* (SVM) [10], *Naïve Bayes Classifier* (NBC) [11], C 4.5 [12], dan *K-Nearest Neighbors* [13].

Penelitian mengenai klasifikasi *cyberbullying* pada komentar Instagram sebelumnya sudah pernah dilakukan pada penelitian Wanda, dkk [10]. Pada penelitian tersebut nilai sentimen dari setiap komentar diklasifikasi ke dalam dua kelas yaitu positif dan negatif *cyberbullying*. Algoritma klasifikasi yang digunakan adalah SVM dengan mengimplementasikan fitur *Lexicon Based Features*. Berdasarkan hasil pengujian diperoleh nilai akurasi terbaik sebesar 90% tanpa fitur *Lexicon Based Features* dan 87% dengan fitur *Lexicon Based Features*. Penelitian klasifikasi teks untuk mendeteksi *cyberbullying* juga pernah dilakukan oleh Kelly Reynolds, dkk [12]. Algoritma yang digunakan untuk klasifikasi adalah algoritma C4.5. Dari hasil pengujian akurasi yang didapatkan adalah sebesar 78.50%.

Kemudian penelitian klasifikasi teks lainnya dilakukan oleh Muljono, dkk dengan menggunakan algoritma NBC [11]. Klasifikasi teks pada penelitian tersebut dilakukan untuk mengetahui opini masyarakat tentang situs belanja *online* pada Twitter yang digunakan untuk menentukan *rating* situs belanja *online*. Hasil dari pengujian menunjukkan bahwa NBC mampu melakukan klasifikasi *tweet* ke dalam dua kelas yaitu positif dan negatif secara akurat dengan rata-rata akurasi sebesar 93.33%.

Penelitian klasifikasi teks dengan menggunakan algoritma *K-Nearest Neighbor* dilakukan oleh Winda, dkk [13]. Klasifikasi dilakukan untuk mendapatkan opini masyarakat terhadap tayangan televisi pada Twitter. Dari pengujian didapatkan hasil akurasi tertinggi sebesar 82,50% menggunakan pembobotan tekstual. Kemudian hasil akurasi terendah sebesar 60% dengan tanpa pembobotan tekstual. Kombinasi dari keduanya menghasilkan akurasi tertinggi sebesar 83.33%.

Dari berbagai referensi penelitian yang telah dilakukan baik algoritma SVM dan NBC sama-sama memiliki akurasi yang tinggi dan cocok digunakan untuk melakukan pengklasifikasian. Berdasarkan latar belakang di atas, maka penelitian ini akan menggunakan kedua algoritma tersebut yaitu SVM dan NBC untuk mengklasifikasikan komentar pengguna Instagram ke dalam dua kelas yaitu *cyberbullying* dan bukan *cyberbullying*. Hasil analisis dari penelitian ini akan ditampilkan dalam bentuk aplikasi berbasis *website* yang membutuhkan waktu singkat untuk pemrosesan halaman. Oleh karena itu, selain membandingkan hasil akurasi, penelitian ini juga akan membandingkan kedua algoritma tersebut dalam hal performansi yang dibutuhkan saat proses klasifikasi dan prediksi.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas maka rumusan masalah dalam penelitian ini adalah menentukan algoritma dengan akurasi dan performansi terbaik untuk melakukan klasifikasi komentar pengguna di Instagram ke dalam kelas *cyberbullying* dan bukan *cyberbullying*.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah menganalisis tingkat akurasi dan performansi algoritma *Support Vector Machine* dan *Naïve Bayes Classifier* dalam mengklasifikasikan komentar pengguna di Instagram ke dalam kelas *cyberbullying* dan bukan *cyberbullying*.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari adanya penelitian ini yaitu :

1. Mengetahui seberapa besar tingkat akurasi dan performansi algoritma *Support Vector Machine* dan *Naïve Bayes Classifier*.
2. Mengetahui algoritma yang sesuai untuk klasifikasi komentar pengguna di Instagram ke dalam kelas *cyberbullying* dan bukan *cyberbullying*.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Penelitian ini hanya melakukan klasifikasi terhadap komentar Bahasa Indonesia.
2. Data yang dikumpulkan untuk dijadikan *sample* adalah sebanyak 1100 komentar pengguna Instagram.
3. Pengumpulan data dilakukan secara *realtime* selama layanan Instagram API masih didukung oleh pihak Instagram.
4. *Emoticons* yang terdapat pada komentar akan diabaikan ketika proses analisis.