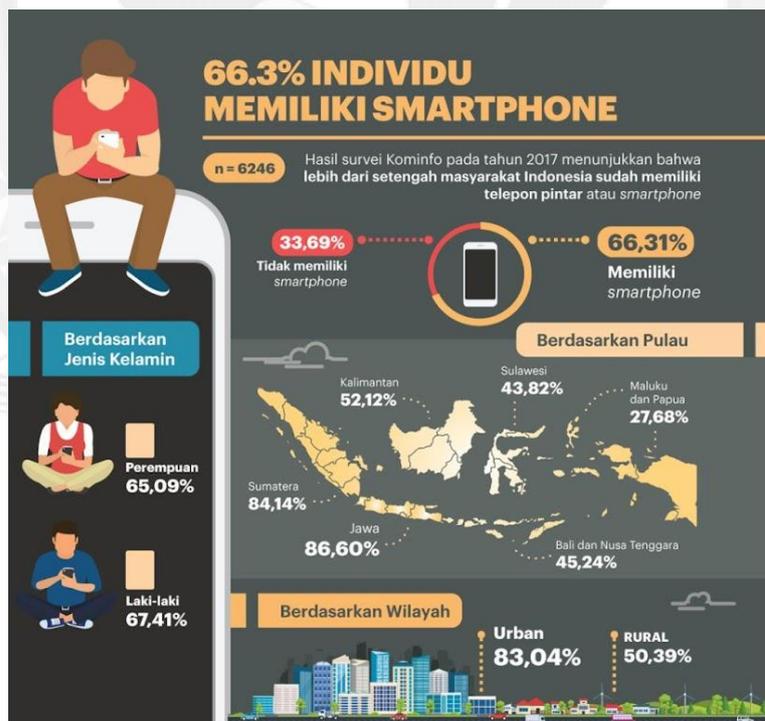


# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi sudah berkembang sangat pesat. Hampir setiap orang memiliki ponsel pintar atau yang biasa disebut dengan *smartphone*. Menurut survei Kominfo pada tahun 2017 menunjukkan bahwa 66,3% masyarakat di Indonesia memiliki *smartphone*. Selain digunakan untuk menerima panggilan dan membalas pesan, *smartphone* juga digunakan sebagai alat untuk browsing, membaca berita hingga menggunakan sosial media. Sosial media merupakan *platform* di mana semua orang dapat mengekspresikan dirinya melalui unggahan status, foto, ataupun video. Selain mengunggah, pengguna juga dapat memberikan komentar terhadap unggahan seseorang dalam sosial media tersebut. Seperti media sosial YouTube yang dimana semua orang dapat bebas berkomentar pada unggahan video yang ada.



Gambar 1. 1 Survey Pengguna *Smartphone*

Sumber: <https://indonesiabaik.id/infografis/663-masyarakat-indonesia-memiliki-smartphone-8>

Akibat dari kebebasan berkomentar yang dapat dilakukan pada media sosial tersebut, orang-orang menjadi kurang mengerti batasan apa saja yang perlu dilakukan saat memberikan komentar. Bukan hanya komentar-komentar positif yang diterima, tetapi tidak jarang ditemukan komentar-komentar yang merujuk ke arah ujaran kebencian ( *hate speech* ) (Suciartini & Sumartini, 2018). Tetapi masih sedikit dari kita yang dapat membedakan mana yang dapat dikategorikan dengan ujaran kebencian dan mana yang bukan. Terkadang pengguna menyalahartikan pernyataan yang disampaikan seseorang sebagai ujaran kebencian padahal pernyataan tersebut hanyalah sebuah kritikan.

Ujaran kebencian yang diterima seseorang berdampak besar pada pola pikir maupun sikap seseorang (Mawarti, 2018). Dengan mengimplementasikan algoritma TF-IDF dan *Support Vector Machine*, kita dapat mengindikasikan apakah komentar yang disampaikan oleh pengguna dapat dikategorikan sebagai ujaran kebencian atau bukan. Dengan begitu, perselisihan antara dua belah pihak dapat diselesaikan jika terciptanya indikator ini.

Alasan penulis melakukan penelitian ini adalah agar dapat membantu orang untuk lebih mudah membedakan mana komentar yang dikategorikan sebagai ujaran kebencian dan yang hanya sekedar merupakan kritikan biasa.

## **1.2 Rumusan Masalah**

Adapun rumusan masalah yang dibahas dalam penelitian ini mencakup :

1. Bagaimana cara mengaplikasikan *web scraping tools* pada kolom komentar akun YouTube Cretivox.
2. Bagaimana cara menerapkan metode TF-IDF dan *Support Vector Machine* ke dalam kalimat komentar pada akun YouTube Cretivox.
3. Bagaimana tingkat akurasi yang diberikan oleh algoritma TF-IDF dan *Support Vector Machine* pada pendeteksian ujaran kebencian.

### 1.3 Batasan Masalah

Adapun batasan masalah yang dibahas dalam penelitian ini mencakup :

1. Data *input* yang digunakan pada penelitian ini adalah kalimat komentar pada kolom komentar video unggahan Cretivox.
2. Jumlah komentar yang digunakan berjumlah 10.000 komentar.
3. Bahasa yang akan digunakan untuk pengujian adalah bahasa Indonesia.
4. Data yang diambil tidak melibatkan emoji/emoticon.
5. Data *output* yang dihasilkan berupa prediksi terhadap data yang di-*input* terdeteksi ujaran kebencian atau tidak.

### 1.4 Tujuan Penelitian

Adapun tujuan penelitian yang dibahas dalam penelitian ini mencakup :

1. Mengaplikasikan *web scraping tools* untuk mengumpulkan komentar yang terdapat dalam kolom komentar akun YouTube Cretivox.
2. Menerapkan metode TF-IDF dan *Support Vector Machine* ke dalam kalimat komentar pada akun YouTube Cretivox.
3. Mengetahui tingkat akurasi dari hasil perhitungan algoritma TF-IDF dan *Support Vector Machine* untuk pendeteksian ujaran kebencian pada komentar akun YouTube Cretivox.

### 1.5 Manfaat Penelitian

Adapun manfaat penelitian yang dibahas dalam penelitian ini mencakup :

1. Penulis dapat menerapkan ilmu yang didapat selama perkuliahannya sehingga dapat membantu masyarakat yang membutuhkan metode ini.
2. Agar dapat mendeteksi apakah komentar pada akun YouTube Cretivox terdapat ujaran kebencian atau tidak dengan mengimplementasikan metode TF-IDF dan *Support Vector Machine*.
3. Memberikan sebuah pengetahuan kepada masyarakat luas agar dapat membedakan kritikan dan ujaran kebencian.

## 1.6 Batasan Masalah

Laporan tugas akhir ini ditulis dengan menggunakan sistematika penulisan yang dapat dijelaskan sebagai berikut.

### **BAB I            PENDAHULUAN**

Berisi latar belakang pemilihan judul, perumusan masalah, tujuan, batasan masalah, manfaat dan sistematika penulisan.

### **BAB II           LANDASAN TEORI**

Berisi dasar-dasar teori yang digunakan untuk acuan dalam merancang dan mengembangkan tugas akhir meliputi *Term Frequency – Inverse Document Frequency* (TF-IDF) dan *Support Vector Machine* (SVM).

### **BAB III          METODOLOGI**

Berisi tentang penjelasan mengenai rancangan dari sistem pendeteksian yaitu pengumpulan data, *cleansing* dan data *preprocessing* dan proses pengerjaan.

### **BAB IV          HASIL DAN PEMBAHASAN**

Berisi tentang hasil penelitian dan analisa terhadap hasil yang telah dihasilkan dalam proses pembuatan aplikasi yang meliputi implementasi rancangan sistem dan analisa hasil.

### **BAB V           KESIMPULAN DAN SARAN**

Berisi tentang kesimpulan yang dapat ditarik setelah menyelesaikan penelitian pendeteksian ujaran kebencian pada kolom komentar YouTube yang telah dilakukan dan saran untuk pengembangan penelitian selanjutnya.