

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Di era perkembangan teknologi informasi dan komunikasi yang sangat pesat saat ini, media sosial menjadi salah satu tempat bagi masyarakat untuk berbagi informasi. Salah satu media sosial yang digunakan oleh masyarakat yaitu youtube. Youtube merupakan media sosial berbasis video. Ada berbagai jenis konten video yang diunggah oleh para penggunanya. Mulai dari konten musik, edukasi, vlog, trailer film, dan lainnya. Oleh karena itu tidak diragukan lagi bahwa Youtube adalah situs pertama yang akan diakses ketika seseorang membutuhkan video tertentu, tidak terkecuali review *smartphone*. Pada saat ini ketika ingin membeli sesuatu salah satunya yaitu *smartphone* kita seringkali mencari referensi dengan cara melihat review-review dari para *content creator* youtube. Dikarenakan pada masa sekarang banyak pengguna *smartphone* yang belum mengetahui seperti apa produk yang akan dibelinya.

Ulasan atau komentar dari pengguna sangat berpengaruh terhadap niat membeli, dan juga menjadi sumber informasi bagi konsumen terkait dengan kualitas *smartphone* tersebut. Oleh karena itu penelitian ini dilakukan untuk bisa menganalisis opini publik tersebut, apakah opini tersebut bernilai negatif atau positif. Dengan begitu masyarakat dapat mengetahui apakah produk yang di review tersebut bernilai negatif maupun positif dari hasil analisis sentimen tersebut. Untuk mengatasi permasalahan tersebut dilakukan penelitian analisis sentimen melalui data komentar pengguna youtube di *channel* GadgetIn.

Analisis sentimen bagian dari penggalian opini, adalah proses otomatis untuk memahami, mengekstraksi, dan memproses data tekstual untuk mendapatkan informasi. Analisis sentimen dalam penelitian ini adalah proses pengklasifikasian teks menjadi dua kelas, yaitu kelas sentimen positif dan negatif (Muhamad Yusuf Nur & Diaz D. Santika, 2011). Ada banyak teknik klasifikasi yang dapat digunakan salah satunya *Rocchio*. *Rocchio* adalah metode klasifikasi yang membandingkan kesamaan isi data latih dan data uji dengan mempresentasikan semua data vektor (Rocchio dkk., 2021). Klasifikasi *Rocchio* memiliki 6 cara untuk mengkalkulasi

batas kelas dengan memanfaatkan centroid untuk menentukan batas antar kelas. Oleh karena itu, klasifikasi *Rocchio* dipilih sebab memiliki keunggulan dengan tidak adanya *underflow* seperti pada klasifikasi Naïve Bayes (Riana & Sibaron, 2021).

Sebagai pendukung proses olah data, terdapat beberapa penelitian sebelumnya yang dilakukan oleh (Ashari dkk., 2020) Perbandingan Kinerja *Algoritma Multinomial Naïve Bayes (MNB)*, *Multivariate Bernoulli* Dan *Rocchio Algorithm* Dalam Klasifikasi Konten Berita Hoax Berbahasa Indonesia Pada Media Sosial menunjukkan hasil dari algoritma *Rocchio* yaitu akurasi sebesar 76%, presisi 88,24% dan *recall* 60%. Peneliti selanjutnya (Rocchio dkk., 2021) Penerapan *Rocchio Classification* dalam Klasifikasi Testimoni Pelanggan Taksi menunjukkan hasil dari algoritma *Rocchio* yaitu akurasi 80%, rata-rata presisi 83,40%, rata-rata *recall* 63,28%, dan rata-rata *f-measure* 65,16%. Dan pada penelitian selanjutnya yang dilakukan oleh (Harpizon dkk., 2022) yang berjudul Analisis Sentimen Komentar Di Youtube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naïve Bayes menunjukkan hasil dari algoritma Naïve Bayes yaitu akurasi 87%, presisi 91%, dan *recall* 97%.

Berdasarkan paparan di atas, penulis akan melakukan penelitian yang berjudul “Analisis Sentimen Pada Komentar Video Ulasan *Smartphone* di Youtube Menggunakan Metode *Rocchio*”. Penggunaan *Rocchio* dapat diharapkan menghasilkan nilai akurasi, presisi, dan *recall* yang tinggi.

## 1.2. Rumusan Masalah

Dari latar belakang yang sudah dijelaskan, maka diambil perumusan masalah dalam penelitian yaitu:

1. Berapakah presentase tingkat akurasi, presisi, dan *recall* yang didapatkan dari hasil perhitungan *Rocchio Algorithm* dalam mengklasifikasi data komentar pada video ulasan *smartphone* di YouTube?
2. Berapakah sentimen negatif dan positif yang diperoleh dari data komentar pada video ulasan *smartphone* di YouTube?

### 1.3. Tujuan Penelitian

Dari rumusan masalah yang disusun akan diperoleh hasil sebagai tujuan dalam penelitian ini yaitu:

1. Mengetahui presentase nilai akurasi, presisi, dan recall yang diperoleh dari metode *Rocchio* dalam mengklasifikasi data komentar pada video ulasan *smartphone* di YouTube.
2. Mengetahui presentase nilai sentimen negatif dan positif yang diperoleh dari data komentar video ulasan *smartphone* di YouTube.

### 1.4. Manfaat Penelitian

Adapun manfaat yang diperoleh dari penelitian adalah:

1. Dapat memberikan kontribusi secara keilmuan berupa analisis sentimen pada video ulasan *smartphone* di YouTube menggunakan metode *Rocchio*.
2. Dapat digunakan sebagai referensi untuk penelitian-penelitian selanjutnya dibidang sentimen.

### 1.5. Batasan Masalah

Dalam penelitian ini penulis memberikan batasan untuk menghindari bias dan pelebaran topik penelitian. Berikut batasan-batasan dalam penelitian ini:

1. Data yang digunakan adalah data komentar pada video ulasan *smartphone* di channel YouTube GadgetIn yang berjudul “Ini soal PRIORITAS – Review realme 10 Indonesia!”.
2. Total data yang digunakan yaitu 500 komentar.
3. Menggunakan *K-Fold* 2, 4, 5 dan 10.