

BAB I PENDAHULUAN

1.1 Latar Belakang

Karena perkembangan teknologi saat ini, kini kita dapat menikmati konten dalam bentuk digital yang dapat kita akses melalui aplikasi dan internet. Salah satu konten audio – visual berupa aplikasi salah satunya adalah YouTube [1]. Oleh karena itu, media sosial YouTube dapat menjadi wadah bagi para masyarakat untuk menunjukkan karya mereka. Karena hal ini, akhirnya banyak orang yang memiliki keinginan untuk mengunggah karya mereka di platform YouTube [2].

Pada platform YouTube, beberapa konten *creator* YouTube mengunggah video yang berisi pengalaman makan mereka di sebuah café yang bernama Karen's Diner. Pada video tersebut, terlihat bahwa anda akan disambut dan diberikan pelayanan yang kasar dan tidak sopan dari café Karen's Diner. Namun hal ini memang menjadi ciri khas dari café Karen's Diner. Di dalam website resmi Karen's Diner juga terdapat kata – kata *we hate good service* dimana jika di terjemahkan dalam Bahasa Indonesia adalah kami benci pelayanan yang baik. Namun untuk cita rasa makanan pada café Karen's Diner juga tidak mereka lupakan, mereka tetap berusaha untuk memberikan makanan yang dapat dinikmati oleh pelanggan café Karen's Diner [3]. Dari video tersebut terdapat banyak komentar opini atau sentiment yang beragam terhadap café Karen's Diner. Oleh karena itu penulis dapat menggali kesimpulan berupa opini – opini masyarakat terhadap pandangan mereka melalui komentar youtube [4].

Karena kepuasan pelanggan tentunya dapat mempengaruhi kelangsungan suatu perusahaan, khususnya industri makanan dan minuman. Kualitas dalam pelayanan adalah berupa interaksi antara pelanggan dan karyawan dengan potensi seperti kemampuan yang profesional, menguasai ilmu dan pengetahuan, memiliki rasa kepedulian kepada pelanggan, dan tentunya pelayan harus dapat memiliki kemampuan untuk menangani serta melayani pelanggan agar dapat tertuju kepada kepuasan pelanggan [5]. Pada SERVQUAL (*Service Quality*) yang dikembangkan

oleh Pasuraman et. al. juga disebutkan bahwa sopan santun para karyawan dapat membangkitkan rasa kepercayaan konsumen yang merupakan suatu alat ukur terhadap kualitas pelayanan [6].

Untuk saat ini penulis belum menemukan penelitian yang membahas mengenai café Karen's Diner khususnya dalam bidang sains dan teknologi. Alternative untuk mengetahui makna dari setiap kata dalam komentar youtube tersebut dengan menggunakan text mining . Text mining merupakan sebuah teknologi kecerdasan buatan yang menggunakan bahasa alami NLP (*natural language processing*) untuk mengubah teks yang tidak beraturan dalam sebuah dokumen menjadi beraturan yang dinormalisasikan untuk kepentingan analisis *machine learning* [7]. Mesin ini memungkinkan manusia dalam memenuhi kebutuhan berbagai hal salah satunya komputasi. Dan *machine learning* adalah salah satunya. *Machine learning* didefinisikan sebagai bidang studi yang memberikan komputer keahlian yang dapat belajar tanpa diprogram secara eksplisit. Ada beberapa jenis algoritma yang dapat digunakan, tergantung kepada masalah yang ingin diselesaikan [8].

Dalam penelitian ini penulis menggunakan tiga metode yaitu *Naïve Bayes*, *K-Nearest Neighbor* dan *Support Vector Mechine*. Pada metode *Naïve Bayes* termasuk pada penambangan data yang populer di sebabkan kemudahan dalam penggunaannya, kemudahan implementasi dengan strukturnya yang sederhana, serta kecepatan pada waktu proses, dan efektivitasnya yang tinggi. Konsep dari *Naïve Bayes* terbentuk oleh prediksi peluang di masa mendatang berdasarkan pengalaman yang terjadi di masa sebelumnya [2]. *K-Nearest Neighbor* merupakan algoritma yang cukup bagus untuk klasifikasi data teks, Metode ini merupakan metode yang sederhana untuk pengklasifikasian tanpa harus melakukan perhitungan secara kompleks, oleh karena itu *K-Nearest Neighbor* juga bisa disebut *lazy learning* [9]. Dan untuk *Support Vector Mechine* merupakan metode klasifikasi yang memprediksi kelas berdasarkan pola dari hasil *training*. Memisahkan kelas positif dan negatif dengan klasifikasi yang dilakukan dengan garis pembatas (*hyperlane*) [10].

Penelitian mengenai analisis sentiment terhadap pandemi Covid-19 terhadap media sosial Twitter di Arab Saudi yang dilakukan oleh Alhaji *et al* mengambil kesimpulan jika *Naïve Bayes* mampu untuk proses klasifikasi teks berupa hasil akurasi yang cukup tinggi yakni sebesar 89%. Tetapi pada penelitian Fithriasari, Mayasari, Irawan, & Winahju memberikan kesimpulan pada penelitian mereka bahwa algoritma *Support Vector Machine* memiliki nilai AUC yang lebih baik dibandingkan dengan *Naïve Bayes* terhadap studi kasus evaluasi pemerintah Surabaya pada media sosial Twitter. Sementara itu jurnal pada penelitian Huq, Ali, & Rahman berupa pengklasifikasian data tweet memberikan kesimpulan jika algoritma *K-Nearest Neighbor* memiliki tingkat akurasi lebih tinggi daripada algoritma *Support Vector Machine*. Berdasarkan ulasan tersebut peneliti berencana menggunakan ketiga metode tersebut untuk dibandingkan hasil kinerjanya pada data yang dimiliki oleh penulis [11].

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah di atas, dapat dirumuskan sebuah permasalahan yaitu:

1. Bagaimana cara kerja algoritma *Naïve Bayes*, *K-Nearest Neighbor*, dan *Support Vector Mechine* terhadap sistem analisis komentar pada YouTube?
2. Bagaimana perbandingan kinerja antara algoritma *Naïve Bayes*, *K-Nearest Neighbor* dan *Support Vector Mechine*?

1.3 Batasan Masalah

Untuk mempersempit pembahasan dan memperkecil terjadinya penyimpangan saat proses penelitian. Maka akan dibuat batasan masalah sebagai berikut:

1. Penelitian ini berfokus pada pembahasan analisis sentiment menggunakan algoritma *Machine Learning*
2. Pada penelitian ini membagi kelas analisis menjadi dua yaitu komentar positif dan negatif

3. Data pada penelitian ini diambil dari komentar YouTube pada rentang waktu pada tanggal 16 april 2023 dan hanya berbahasa Indonesia
4. Penelitian ini menggunakan tiga algoritma Machine Learning yaitu Naïve Bayes, Support Vector Machine, dan K-Nearest Neighbor

1.4 Tujuan Penelitian

1. Penelitian ini memiliki tujuan dalam mengklasifikasi informasi dari komentar YouTube, yang nantinya komentar tersebut dapat memiliki sentimen positif ataupun negatif.
2. Pada penelitian ini, penulis akan menguji dan membandingkan mengenai performa ketiga algoritma dengan mempergunakan metrik evaluasi.

1.5 Manfaat Penelitian

Untuk penelitian ini diharapkan dapat menambah wawasan dan pengetahuan mengenai analisis sentiment terhadap dataset yang di ambil melalui komentar YouTube. Untuk pengguna YouTube penelitian ini dapat memberikan informasi mengenai penilaian sebuah konten video berdasarkan ekstrasi dan penambangan data dari komentar dan opini masyarakat di YouTube. Lalu bagi penulis dengan adanya penelitian ini dapat memberikan wawasan dan kemampuan dalam pemecahan masalah text mining dalam analisis sentiment dalam YouTube.

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini sebagai berikut:

BAB I PENDAHULUAN

Bab 1 berisikan mengenai latar belakang bagaimana bisa dilakukannya penelitian ini, rumusan masalah juga akan dibahas pada bab 1, dan juga batasan masalah agar menghindari adanya perluasan pokok masalah supaya penelitian lebih terarah, setelah itu ada tujuan dari penelitian ini dibuat, adapun penelitian dibuat agar memiliki manfaat yang akan dibahas pada bab 1, setelah itu ada sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini berisi mengenai studi literature yang akan mengkaji hasil dari penelitian terdahulu yang masih dalam pembahasan penelitian ini, setelah itu terdapat dasar teori.

BAB III METODE PENELITIAN

Eksperimen penelitian akan dibahas pada bab ini yang berisi mengenai sasaran penelitian, setelah itu ada alur penelitian yang berisi mengenai langkah – langkah penelitian yang di perlihatkan berupa gambar bagan alur penelitian, adapun alat dan bahan juga akan dibahas pada bab 3.

BAB IV HASIL DAN PEMBAHASAN

Pada bab 4 akan membahas mengenai rancangan eksperimen, dan pengerjaan dari eksperimen.

BAB V PENUTUP

Kesimpulan akan dibahas pada bab 5, serta saran.