

BAB I

PENDAHULUAN

1.1. Latar Belakang

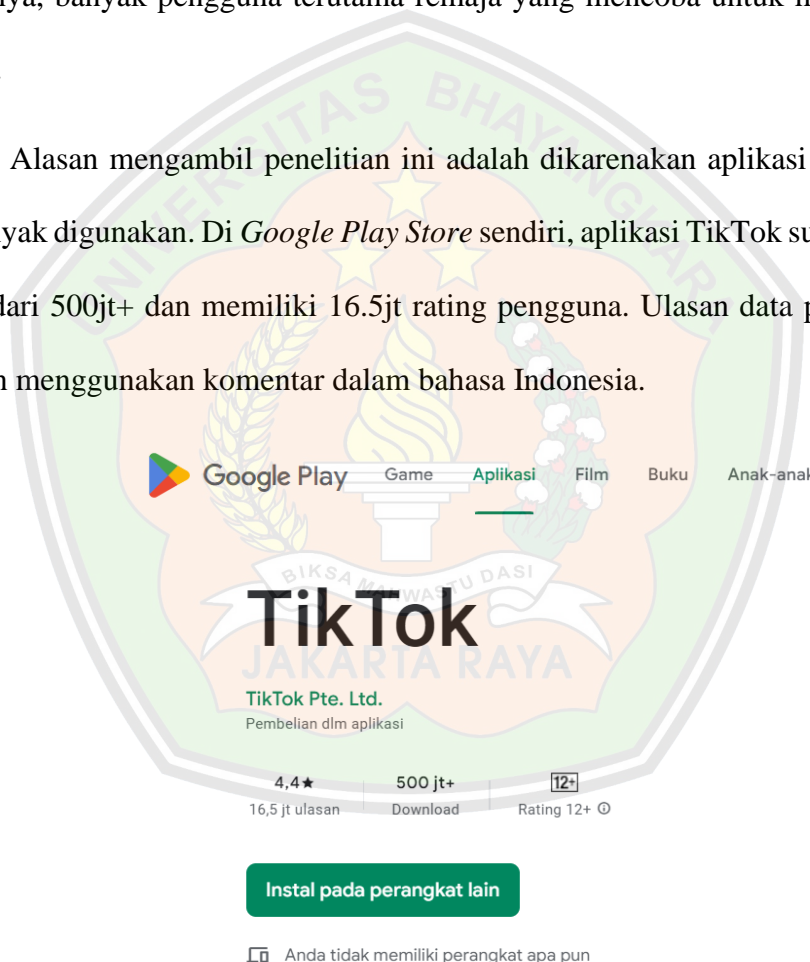
Teknologi sekarang sudah sangat maju, banyak media yang bisa digunakan untuk melakukan komunikasi, media ini sangat mudah digunakan dengan menghubungkan ke internet. Sangat ini, berkat perkembangan teknologi yang sangat pesat, masyarakat sangat mudah untuk berkomunikasi menggunakan salah satu teknologi yang modern yaitu smartphone. Smartphone ini sangat memungkinkan masyarakat untuk mengakses media dengan sangat mudah dan masyarakat dapat menggunakannya untuk berkomunikasi, seperti melalui Facebook, Instagram, Twitter, dll. Media ini memiliki berbagai fungsi. Salah satu aplikasi yang saat ini sedang banyak mendapat perhatian adalah aplikasi TikTok. Aplikasi TikTok ini memungkinkan pengguna membuat video berdurasi pendek dan cepat.

TikTok adalah aplikasi populer di kalangan anak muda. TikTok merupakan aplikasi yang pertama kali diluncurkan di China sebelum hadir di Indonesia pada akhir tahun 2017. Aplikasi TikTok ini resmi diluncurkan oleh Zhang Yiminy pada tahun 2016. Aplikasi ini dapat menambahkan banyak fitur seperti, Menambahkan musik dalam video, mengubah suara, filter, tambahan efek dan stiker, dll. Aplikasi TikTok ini juga mendorong para penggunanya untuk berkreasi saat membuat video. Sementara itu, Aplikasi ini juga sekarang menjadi aplikasi yang populer di kalangan masyarakat, bahkan di masa pandemi jumlah pengguna TikTok meningkat 20% dibandingkan dengan biasanya. Rata-rata usia pengguna TikTok ini berusia hampir

sama dengan rata-rata usia anak di bawah umur dan bahkan kebanyakan orang dewasa [1].

Sayangnya, popularitas TikTok tidak hanya berdampak positif tetapi juga berdampak negatif karena kurangnya citra diri individu tersebut, misalnya mengenakan pakaian seksi dan menari dengan gerakan erotis dan tidak pantas. Ironisnya, banyak pengguna terutama remaja yang mencoba untuk mengikuti hal ini [2].

Alasan mengambil penelitian ini adalah dikarenakan aplikasi TikTok saat ini banyak digunakan. Di *Google Play Store* sendiri, aplikasi TikTok sudah diunduh lebih dari 500jt+ dan memiliki 16.5jt rating pengguna. Ulasan data penelitian ini dengan menggunakan komentar dalam bahasa Indonesia.



Gambar 1.1 Data Ulasan dan Jumlah Unduhan di *Google Play Store*

Pada penelitian ini penulis akan membandingkan antara dua algoritma yaitu *Naïve Bayes* dan C4.5. Penelitian berjudul Analisis sentimen Tanggapan Masyarakat Aplikasi TikTok Menggunakan Metode *Naïve Bayes* dan *Categorical*

Propotional Difference (CPD) yang dilakukan oleh Junda Alfiah Zulqornain, Indriati, Putra Pandu Adikara, 2021 [1] telah berhasil melakukan proses klasifikasi dengan algoritma *Naïve Bayes* dengan akurasi sebesar 0,729947.

Selanjutnya, penelitian juga pernah dilakukan oleh Billy Gunawan, Helen Sasty Pratiwi, Enda Esyudha Pratama, 2018 [3] yang berjudul Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode *Naïve Bayes* dengan akurasi sebesar 77.78%. Selanjutnya, penelitian lain juga dilakukan oleh Dedi Darwis, Nery Siskawati, Zaenal Abidin, 2021 [4] yang berjudul Penerapan Algoritma *Naïve Bayes* untuk Analisis Sentimen *Review* Data Twitter BMKG Nasional dengan akurasi sebesar 69.97%.

Pada penelitian algoritma C4.5 penelitian juga dilakukan oleh Siska Febriani, Heni Sulistiani, 2021 [5] yang berjudul Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5 dengan akurasi sebesar 72.67%. Selanjutnya, penelitian lain juga dilakukan oleh Endang Sri Palupi, Said Mirza Pahlevi, 2020 [6] yang berjudul Klasifikasi *Opportunity* Menggunakan Algoritma C4.5 dan *Naïve Bayes* Berbasis *Particle Swarm Optimazation* dengan akurasi Algoritma C4.5 sebesar 66,67% dan Algoritma *Naïve Bayes* sebesar 75,67%

Beberapa dari penelitian sebelumnya menjadi dasar untuk riset yang akan dilakukan terkait topik analisis sentimen menggunakan algoritma *Naïve Bayes* dan C4.5. Hal ini didasari oleh banyaknya komentar pengguna TikTok yang bersifat positif dan negatif. Komentar yang diperoleh dari ulasan tersebut akan melalui *preprocessing*.

Berdasarkan penelitian terdahulu yang peneliti paparkan di atas, maka penelitian yang peneliti ambil akan merujuk pada penelitian [1]. Pada penelitian [1] peneliti terdahulu menggunakan metode *Naïve Bayes*. Mereka menyatakan *Naïve Bayes* semakin besar *term* yang digunakan tidak selalu memberikan hasil yang baik. Hal ini terlihat pada hasil pengujian menggunakan *5-fold cross validation* yang menunjukkan dari nilai akurasi tertinggi, tapi bisa menghasilkan nilai *f-measure* yang rendah.

Pada penelitian [6] terdahulu menggunakan metode *Naïve Bayes* dan C4.5. Mereka menyatakan penelitian ini dilakukan pengujian model dengan membandingkan tiga algoritma C4.5, algoritma C4.5 berbasis *Particle Swarm Optimazation*, algoritma *Naïve Bayes*. Hasil evaluasi dan validasi, diketahui bahwa algoritma C4.5 berbasis *Particle Swarm Optimazation* memiliki nilai akurasi paling tinggi sebesar 80.90% dibandingkan dengan algoritma *Naïve Bayes* yang ada pada urutan kedua akurasinya yaitu sebesar 75.67% dan paling rendah akurasinya algoritma C4.5 dengan akurasi sebesar 66.67%. Dengan demikian, metode C4.5 berbasis *Particle Swarm Optimazation* merupakan metode yang cukup baik dalam memprediksi *oppurtinity* yang akan *Purchase Order* (Buy). Berangkat dari penelitian mereka inilah, maka peneliti akan melakukan penelitian serupa, namun yang membedakannya adalah penelitian ini diharapkan akan bisa lebih baik dari hasil akurasi sebelumnya.

Berdasarkan latar belakang permasalahan di atas, maka penelitian yang akan dilakukan akan mengambil judul **“Implementasi Algoritma *Naïve Bayes* dan**

Algoritma C4.5 Untuk Melakukan Analisis Sentimen Terhadap Ulasan Komentar Pengguna TikTok di *Google Play Store*”

1.2. Identifikasi Masalah

Berdasarkan latar belakang di atas, maka masalah yang teridentifikasi sebagai berikut:

1. Belum adanya teknik klasifikasi untuk mengelompokkan ulasan atau *review* positif dan negatif pada TikTok di *Google Play Store*.
2. Belum adanya penerapan algoritma *Naïve Bayes* dan C4.5 dalam ulasan TikTok di *Google Play Store*.

1.3. Batasan Masalah

Adapun batasan masalah yang digunakan peneliti agar pembahasan dalam penelitian ini tidak menyimpang adalah:

1. Penelitian ini menggunakan data ulasan komentar pengguna pada aplikasi TikTok di *Google Play Store*.
2. Data yang diambil hanya ulasan yang berbahasa Indonesia.
3. Kategori yang digunakan adalah positif dan negatif.
4. Data yang diambil dari bulan April 2023 – Juni 2023.

1.4. Rumusan Masalah

Berdasarkan penjelasan yang telah diuraikan, maka dapat ditentukan rumusan masalah sebagai berikut:

1. Bagaimana tingkat akurasi algoritma *Naïve Bayes* dan C4.5 dengan seleksi fitur TF-IDF?
2. Bagaimana mengimplementasi algoritma *Naïve Bayes* dan C4.5 dengan seleksi fitur TF-IDF?

1.5. Tujuan Penelitian

Berdasarkan identifikasi masalah di atas, tujuan penelitiannya adalah sebagai berikut:

1. Menerapkan algoritma *Naïve Bayes* dan C4.5 dalam pengelompokan ulasan pengguna aplikasi TikTok.
2. Menentukan perbandingan algoritma *Naïve Bayes* dan C4.5 dalam pengelompokan ulasan aplikasi TikTok.

1.6. Manfaat Penelitian

Adapun manfaat yang bisa diambil dari hasil penelitian ini adalah sebagai berikut:

1. Diharapkan bermanfaat bagi pembaca yang ingin mempelajari *Naïve Bayes* dan C4.5
2. Memiliki kepentingan dapat melihat informasi yang ada dalam kumpulan ulasan penggunaan aplikasi TikTok dari berbagai komentar untuk pengguna baru yang ingin menggunakan aplikasi TikTok.

3. Mengetahui ketepatan algoritma *Naïve Bayes* dan C4.5 dalam pengelompokan ulasan aplikasi TikTok.

1.7. Sistematika Tugas Akhir

Tulisan merupakan gambaran dari susunan tulisan itu sendiri yang dibuat secara teratur dan rinci sehingga dapat tersaji gambaran yang utuh. Sistem penulisannya adalah sebagai berikut

BAB I PENDAHULUAN

Bab ini menjelaskan latar belakang, identikasi masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika tugas akhir.

BAB II LANDASAN TEORI

Bab ini menjelaskan teori yang berhubungan dengan pokok pembahasan dalam penelitian dan secara garis besar berisi tentang analisis sentimen.

BAB III METODOLOGI PENELITIAN

Bab ini membahas penjelasan secara bertahap dan terperinci tentang langkah-langkah yang digunakan untuk membuat kerangka berfikir dan kerangka kerja dalam menyelesaikan analisis sentimen menggunakan algoritma *Naïve Bayes* dan C4.5 terhadap aplikasi TikTok.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini menunjukkan hasil dari ulasan yang peneliti lakukan, membahas metode, mengukur hasil akurasi dengan algoritma klasifikasi.

Hasil dari penerapan algoritma *Naïve Bayes* dan C4.5 pada dataset yang di ambil dari *review* aplikasi TikTok.

BAB V PENUTUP

Pada bab ini membahas kesimpulan dan saran dari penelitian.

