

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era digital saat ini, kemajuan teknologi informasi semakin pesat. Selain sebagai sarana komunikasi dan informasi, internet juga memungkinkan individu untuk meningkatkan pendapatan ekonomi dengan mengabdikan keinginan konsumen dalam berbelanja, khususnya melalui belanja online. Belanja online merupakan kebutuhan bagi individu tertentu agar tetap eksis sebagai konsumen. *shopee* adalah aplikasi yang memfasilitasi jual beli online dan mudah diakses melalui ponsel pintar. Pengguna aplikasi *shopee* dapat berbelanja terlebih dahulu hingga batas yang ditentukan menggunakan fungsi *paylater*. *Shopeepaylater* adalah nama fungsi *paylater* *shopee*. *Shopee* memberikan penghargaan kepada pengguna aktif program *shopee* dengan pinjam uang cepat dengan bunga yang sangat rendah melalui fungsi *Shopeepaylater*. Pengguna akan diminta untuk memberikan informasi pribadi, foto diri, dan foto KTP mereka untuk memanfaatkan layanan ini. Setiap pengguna baru yang ingin memanfaatkan fungsi *Shopeepaylater* mungkin penasaran untuk mengetahui bagaimana pengguna lain menanggapi *tweet* mereka menggunakan opsi *Shopeepaylater* di *twitter*. Untuk memastikan seberapa besar kemudahan yang ditawarkan opsi *Shopeepaylater* ini kepada konsumen, diperlukan analisis sentimen. Oleh karena itu kami menginginkan solusi yang dapat menerapkan analisis sentimen terhadap opini *tweet* pengguna *Shopeepaylater* di *twitter*. Analisis sentiment juga dikenal sebagai opinion mining adalah bidang studi yang mencoba mengkaji sikap, pandangan, sentimen, dan penilaian orang terhadap berbagai objek, termasuk barang, jasa, organisasi, isu, masalah, peristiwa, dan tema, seperti serta karakteristik [1].

Tujuan dasar dari analisis sentimen adalah untuk mengklasifikasikan berbagai pernyataan atau pendapat yang positif atau negatif dari sebuah teks yang ada pada kalimat yang berada di dalam dokumen tertentu [2]. Dalam melakukan analisis sentiment, terdapat banyak metode perhitungan probabilitas yang membantu dalam pengklasifikasian, antara lain *Support Vector Machine (SVM)*, Algoritma *Decision Tree C4.5*, *K-Nearest Neighbors (KNN)*, dan *Naïve Bayes*

Classifier (NBC). Dalam penerapannya, sebelum dilakukan klasifikasi perlu dilakukan beberapa tahapan. Data yang mulanya berbentuk teks harus diekstraksi dan dibobotkan terlebih dahulu dengan menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)* yang mana metode ini bertujuan untuk melakukan pembobotan kata dari proses ekstraksi kata dengan menggunakan perhitungan kata umum [3]. Fitur yang dihasilkan dari metode *Term Frequency-Inverse Document Frequency (TF-IDF)* nantinya akan banyak karena data yang digunakan berupa data teks. Dalam penelitian ini memanfaatkan algoritma *Naïve Bayes Classifier (NBC)* yang digunakan untuk mengklasifikasikan sebuah *tweet* kedalam kategori sentimen positif atau negatif serta *K-Fold Cross Validation* sebagai metode untuk evaluasi modelnya. Banyaknya data berupa *tweet* dari *twitter* yang ingin dikategorikan atau diberi label menjadi dasar metode *Naive Bayes* yang digunakan dalam penelitian ini. Metode ini sering diterapkan karena kemudahan dan kesederhanaannya. Metode ini berpotensi baik dan banyak digunakan dalam teknik klasifikasi terutama *twitter* [4].

Pada penelitian yang berjudul “*Perbandingan Algoritma Naïve Bayes Dan K-Nearest Neighbor pada Analisis Sentimen Pembukaan Pariwisata Di Masa Pandemi Covid 19*”. Pada penelitian ini proses algoritma *Naïve Bayes* dibagi menjadi dua, yaitu data training yang digunakan sebagai acuan klasifikasi data testing dan data testing yang digunakan untuk menentukan kategori dengan nilai tertinggi. Sedangkan *K-Nearest Neighbor* memiliki prinsip bekerja berdasarkan jarak terpendek dari sampel uji ke sampel latih. Kedua metode tersebut menggunakan *confusion matrix* sebagai tahapan evaluasi dan diketahui bahwa algoritma *Naïve Bayes Classifier* memiliki nilai akurasi sebesar 75,53%, sedangkan metode *K-Nearest Neighbor* sebesar 48,66% [5]. Pada penelitian yang berjudul “*Perbandingan Algoritma Naïve Bayes, K-NN, dan SVM dalam Pengklasifikasian Sentimen Media Sosial*”. Dari penelitian ini ketiga metode tersebut menggunakan *confusion matrix* sebagai tahapan evaluasi, diketahui bahwa algoritma *Naïve Bayes Classifier* memiliki nilai akurasi lebih tinggi sebesar 79.8% dibandingkan algoritma *K-NN* sebesar 50.23% dan *SVM* sebesar 75.29%, sehingga algoritma *Naïve Bayes Classifier* lebih tepat digunakan untuk mengklasifikasi ulasan aplikasi media sosial [6].

1.2 Perumusan Masalah

1.2.1 Permasalahan

Platform shopee menyediakan opsi pembayaran bernama *Shopeepaylater* yang memungkinkan konsumen untuk membeli barang yang diinginkan dan membayarnya ketika sudah jatuh tempo di kemudian hari. Sekarang, ada banyak pro dan kontra di *twitter* tentang layanan *paylater*. Respon pengguna lama yang telah menggunakan *Shopeepaylater* sebagai semacam testimoni setelah menggunakan fitur ini melalui *twitter* dalam bentuk *tweet* adalah sesuatu yang pasti menarik untuk dipelajari oleh pelanggan baru *Shopeepaylater*. Melalui analisis deskriptif, pengguna *Shopeepaylater* akan memberikan komentar positif dan negatif. Sehingga diperlukan analisis sentimen untuk mengetahui apakah lebih banyak sentimen positif atau sentimen negatif dari masyarakat.

1.2.2 Metode Usulan

Penelitian ini menggunakan metode *Naïve Bayes Classifier* sebagai metode pengklasifikasian serta *K-Fold Cross Validation* sebagai metode untuk evaluasi modelnya.

1.2.3 Pertanyaan Penelitian

1. Bagaimana hasil klasifikasi kategori sentimen dari pengguna *Shopeepaylater* dengan menggunakan *Naïve Bayes Classifier* ?
2. Berapa besar tingkat akurasi yang didapatkan dari metode klasifikasi *Naïve Bayes Classifier* dalam penentuan sentimen pengguna *Shopeepaylater* ?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

1. Mengidentifikasi sentimen dari pengguna *Shopeepaylater* berdasarkan *tweet* pada *twitter* dengan menggunakan *Naïve Bayes Classifier*.
2. Menghitung tingkat akurasi yang didapatkan dari analisis sentimen dengan menggunakan *Naïve Bayes Classifier*.

1.3.2 Manfaat

1. Memperoleh berbagai macam informasi dari pendapat layanan *Shopeepaylater* terhadap pelayanan yang telah diberikan sehingga dapat dijadikan sebagai perbandingan saat memilih pinjaman.
2. Mengetahui kemampuan algoritma *Naïve Bayes Classifier* untuk menganalisis sentimen pada *tweet* pengguna *Shopeepaylater* pada media sosial *twitter*.

1.4 Batasan Masalah

Supaya penelitian yang dilakukan tidak menyimpang dan keluar dari pokok permasalahan maka dibutuhkan batasan sehingga penelitian menjadi lebih terfokus. Beberapa batasan dalam penelitian ini adalah sebagai berikut:

1. *Tweet* yang digunakan merupakan *tweet* yang berbahasa Indonesia.
2. Sentimen yang digunakan yaitu berupa sentimen positif (kepuasan) dan sentimen negatif (kritikan) terhadap pelayanan *Shopeepaylater*.
3. Data yang digunakan adalah data yang diambil dari *twitter* sebanyak 1.024 data *tweet* yang diambil per tanggal 1 Oktober 2022 sampai 1 April 2023 yang berkaitan dengan pendapat pelanggan terhadap kualitas *Shopeepaylater*.

1.5 Sistematika Penulisan

BAB I. PENDAHULUAN

Pada bab ini memuat tentang latar belakang penelitian, perumusan masalah, tujuan dan manfaat, batasan masalah, dan sistematika penulisan.

BAB II. TINJAUAN PUSTAKA

Pada bab ini berisi uraian mengenai tinjauan pustaka dan kerangka teori yang relevan dan terkait dengan penelitian yang dilakukan.

BAB III. METODE USULAN

Pada bab ini membahas mengenai rancangan dari sistem yang akan dibuat kedepannya setelah melakukan pengumpulan data dan analisa masalah mulai

dari kebutuhan software, hardware, rancangan atau pola arsitektur serta alur dari sistem yang akan digunakan.

BAB IV. HASIL DAN PEMBAHASAN

Bab ini menjelaskan mengenai hasil dan implementasi dari analisis sentimen tentang pengguna shopeepaylater dengan metode naïve bayes.

BAB V. PENUTUP

Bab ini merumuskan kesimpulan dari hasil penelitian yang dilakukan serta memberikan saran terhadap para pembaca.

DAFTAR PUSTAKA

Memuat daftar jurnal-jurnal atau referensi yang digunakan sebagai acuan dalam penelitian.