

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perbaikan mekanis terbantu pada kenaikan pemanfaatan akses jaringan website di Indonesia. Di Indonesia sendiri, klien website pada tahun 2018 menghubungi 171, 17 juta orang. Perihal ini pula secara positif mengganti metode berperilaku ataupun kecenderungan warga setempat dalam melaksanakan bermacam perihal. Salah satunya merupakan strategi pertukaran yang lebih simpel serta gampang menyesuaikan diri. Salah satu tata cara simpel buat mengeksekusinya merupakan dengan menggunakan aplikasi bisnis berbasis website. Tokopedia, shopee, serta bukalapak jadi panggung bisnis online yang sangat banyak didatangi di Indonesia. Persaingan yang sungguh- sungguh antar pelakon bisnis berbasis website menimbulkan organisasi wajib menjajaki watak administrasi dari tiap- tiap organisasi..(Sasmita et al., n.d.)

Tokopedia ialah toko berbasis internet yang sangat banyak didatangi warga Indonesia dengan jumlah wisatawan 1, 2 miliar, terdiri 1 juta wisatawan dari website portabel serta 8 juta wisatawan dari zona kerja. Walaupun terletak di peringkat 5 paling atas yang tidak henti- hentinya menemukan nilai paling tinggi, tetapi yang jelas tidak seluruh dari mereka membagikan pendapat positif tentang aplikasi Tokopedia, bisa jadi terdapat yang membagikan pendapat negatif, terdapat pula yang membagikan pendapat non- partisan tentang aplikasi ini, komponen klien ataupun keyakinan wilayah setempat memegang peranan berarti dalam melindungi toko di website, dalam sebagian permasalahan benda yang dijual ataupun dibeli lewat salah satu tujuan bisnis berbasis website benda dagangan( benda) tidak cocok dengan gambar di promosi serta Terlebih lagi terkadang produk tidak cocok ataupun apalagi didapat dalam keadaan rusak, tidak cocok dengan kebutuhan.(Apriani et al., 2019)

Dilansir dari Techinasia, Bhayu Rareli Arsyad sebagai Senior Client Experience Tokopedia berkata kalau suara klien lewat hiburan berbasis website ialah sumber informasi yang signifikan serta bisa diolah jadi suatu pengetahuan yang berharga, tidak hanya itu suara klien bisa membagikan usulan kepada Tokopedia serta bisa ditindaklanjuti oleh Tokopedia sehingga bantuannya dapat diterima. Salah satu fokus terpentingnya merupakan opini sosial yang diberikan oleh klien hiburan virtual selaku penghormatan yang bisa membujuk pelanggan baru buat bergeser ke Tokopedia. (Pajri et al., 2020)

Pada halaman Google Play, terdapat fitur komentar yang berfungsi sebagai tempat bagi pengguna untuk berbagi pandangan mengenai aplikasi yang mereka gunakan. Berbagai jenis ulasan ditampilkan di sana, baik yang bersifat negatif maupun positif, dan hal ini merubah cara pandang pada pengguna lainnya. Mengumpulkan dan mengelola data ulasan bukanlah tugas yang mudah karena volume ulasan di platform ini seringkali sangat besar. Moraes et al. (2013) menyatakan bahwa pendekatan yang tepat untuk mengatasi tantangan ini adalah melalui teknik web scraping. Maka karena itu, diperlukan pengklasifikasikan ulasan tersebut menjadi kategori emosi masing-masing ulasan, dengan metode text mining. Terkadang, dalam analisis sentimen, muncul masalah ketidakseimbangan data di mana jumlah ulasan dalam beberapa kelas tidak seimbang. Hal ini dapat memengaruhi kinerja model, terutama dalam mengenali kelas minoritas. Dalam konteks ini, algoritma pembelajaran mesin cenderung menghasilkan model dengan keengganan rendah terhadap kelas minoritas sambil mengelola informasi yang timpang. Beberapa solusi telah diusulkan untuk mengatasi masalah ketidakseimbangan dataset, termasuk klasifikasi berbasis kelas, resampling, dan pembelajaran dengan biaya sensitif. Metode (SMOTE) diterapkan untuk mengatasi masalah ketidakseimbangan data. Metode ini bertujuan untuk menghasilkan sampel sintetis dari kelas minoritas agar dataset menjadi lebih seimbang, yang pada gilirannya akan meningkatkan kinerja algoritma pengklasifikasi dalam analisis sentimen yang diterapkan..(Putri & Kharisudin, 2022)

Penelitian ini mengaitkan prosedur penambangan data yang berencana memikirkan

pengelompokan dalam pengecekan opini dari perspektif klien bisnis berbasis website yang sudah disusun di playstore memakai strategi K- Closest Neighbor. K- Closest Neighbor ialah sesuatu metode buat mengkarakterisasi objek dengan mempersiapkan data yang menggunakan jarak ataupun kemiripan terdekat dengan objek tersebut. Pada sesi pendidikan, perhitungan ini melaksanakan kapasitas vektor elemen serta pengelompokan data pendidikan. Keunggulan metode K- Closest Neighbor merupakan perhitungan urutan yang didemonstrasikan buat menggapai hasil yang sangat presisi serta cocok dengan ditaksir yang diterapkan dalam sesuatu aplikasi. (Firdaus et al., 2022)

## 1.2 Identifikasi Masalah

Masalah yang dikenali di eksplorasi sebagai berikut::

1. Petunjuk langkah demi langkah untuk melakukan Perhitungan K-Nearst Neighbor (K-NN) untuk survei perasaan investigasi pada audit aplikasi Tokopedia di Google Play.
2. Mengkarakterisasi informasi dengan membaca survei dari Google Play di aplikasi Tokopedia membutuhkan waktu yang sangat lama dan tidak cocok untuk klien baru.

## 1.3 Tujuan

Tujuan dari penelitian ini untuk mengetahui presentase yang di dapat pada penyaringan pengkalsifikasian data ulasan negatif dan ulasan positif menggunakan algoritma K-Nearst Neighbor untuk sentimen aplikasi Tokopedia di ulasan rievew google play. Data yang di kumpulkan melalui review di google play pada aplikasi Tokopedia menggunakan metode web scraping di karenakan metode ini memiliki konsistensi syntax dalam sebagai Bahasa pemrograman dan flexible.

## 1.4 Batasan Masalah

Batasan masalah yang perlu dipertimbangkan adalah:

1. Database pada aplikasi dibuat menggunakan MySQL dan menggunakan Bahasa pemrograman PHP.
2. Algoritma yang digunakan untuk penelitian adalah algoritma K-Nearst Neighbor (KNN).
3. Klasifikasi opini menjadi 2 kategori yaitu opini positif dan negative, pada data computer.
4. Format file dataset menggunakan file csv.

## 1.5 Kontribusi

Kontribusi pada penelitian ini yaitu membantu pengembang aplikasi tokopedia dalam menyaring komentar positif dan negatif dengan mengimplementasikan algoritma K-Nearst Neighbor (KNN).