

# BAB 1

## USULAN GAGASAN

### 1.1 Deskripsi Umum Masalah

#### 1.1.1 Latar Belakang Masalah

Perkembangan yang pesat membuat berbagai aktivitas menjadi mudah dilakukan, seperti transaksi online. Tokopedia dan Shopee merupakan *e-commerce* yang telah mendominasi pasar di Indonesia dalam beberapa tahun terakhir dengan menjadi pemuncak peringkat *e-commerce* [1]. Walaupun *review* pada aplikasi memiliki parameter penilaian seperti bintang dengan range 1-5, namun ada juga yang asal memberikan bintang dan tidak sesuai dengan *review*nya [2]. Untuk itu sentimen analisis memiliki banyak manfaat salah satunya untuk mengetahui apakah para pelanggan memiliki tanggapan yang baik atau tidak terhadap produk dan ini bisa menjadi masukan untuk perkembangan bisnis produk tersebut dimasa depan. Karena manfaat ini, banyak bidang yang menggunakan sentimen analisis, salah satunya para penyedia aplikasi di *Google Play Store* guna mengetahui kualitas aplikasi dengan melihat *review-review* yang diberikan terhadap aplikasi tersebut [3].

Untuk penelitian dan analisis ini, metode dan analisis diperlukan untuk membagi komentar pengguna ke dalam beberapa kategori, yang dalam penelitian ini adalah kategori positif dan negatif. Penelitian ini menggunakan metode *Naïve Bayes* untuk menghasilkan kesan positif dan negatif terhadap komentar pengguna di aplikasi Tokopedia yang ada di *Google Play Store* [4]. Selain itu metode *Support Vector Machine (SVM)* digunakan untuk melakukan analisis *review* pengguna pada aplikasi *Google Play Store*. Untuk metode klasifikasi *sentiment analysis* yaitu menggunakan *supervised* dan *unsupervised*. Klasifikasi *supervised* adalah klasifikasi yang memiliki kontrol terhadap informational classes berdasarkan *stopwords* sampel dan adanya kontrol terhadap keakuratan klasifikasi dan sedangkan untuk *unsupervised* adalah untuk meminimalisir kesalahan operator [5]. *TF-IDF (Term Frequency-Inverse Document Frequency)* adalah sebuah metode proses pengelompokan kata dalam hasil *review* yang terdapat dalam *review* aplikasi Tokopedia dan Shopee. Kemudian diklasifikasikan berdasarkan dengan jenis kata yang ada [6].

Maka dari itu untuk mengatasi permasalahan yang telah dijabarkan, dibutuhkannya *sentiment analysis* dengan algoritma *Naïve Bayes* dan *Support Vector Machine* dengan metode *TF-IDF*. Dengan metode *TF-IDF* dapat mengelompokkan kata-kata berdasarkan hasil *review* pada aplikasi Tokopedia dan Shopee, hasil dari klasifikasi akan ditampilkan di website.

## 1.1.2 Analisa Masalah

Dari latar belakang masalah yang telah dijabarkan, didapatkan hasil permasalahan yang dapat dianalisis dari berbagai aspek diantaranya sebagai berikut.

### 1.1.2.1 Aspek Bisnis dan Pengembangan Aplikasi Aspek Manufakturabilitas

Pada saat ini pertumbuhan belanja melalui *e-commerce* terus menunjukkan peningkatan sehingga ini menjadikan sentimen publik terhadap produk atau layanan menjadi penting. Sehingga dibutuhkan proses pemantauan dan pengklasifikasian terhadap sentimen publik tersebut. Tetapi ini bukanlah pekerjaan mudah karena jumlah sentimen sangat banyak jika proses yang dilakukan secara manual. Sehingga peneliti melakukan penelitian untuk menyelesaikan masalah ini menggunakan algoritma klasifikasi *Support Vector Machine*. Kemudian klasifikasi dari sentimen ini dibagi menjadi dua jenis yaitu positif dan negatif [2]. Sehingga analisis tersebut dapat mempengaruhi keputusan bisnis dan pengembangan aplikasi.

### 1.1.2.2 Aspek Ekonomi Digital

Saat ini internet memungkinkan pengguna untuk membuat ulasan secara online di berbagai jenis *platform*. Salah satunya aplikasi *e-commerce* Shopee dan Tokopedia pada *Google Play Store* dimana kelas sentimen positif dan negatif yang terdapat pada ulasan online jelas mencerminkan persepsi pengguna tentang berbagai jenis layanan dan produk yang ada. Selain itu, pelanggan berpotensi yang membaca ulasan online dapat secara signifikan terpengaruh oleh sentimen dari ulasan yang tertera pada kolom ulasan. Hal ini menandakan ulasan yang bersentimen positif ataupun negatif yang ditinggalkan oleh pengguna sangat mempengaruhi pengguna lainnya dalam memilih layanan maupun produk yang dicari [1].

### 1.1.2.3 Aspek Teknologi

Dalam konteks teknologi, permasalahan yang dapat diidentifikasi dari latar belakang tersebut melibatkan penggunaan teknologi untuk menganalisis dan mengelola *review* pengguna pada *platform Google Play Store*, khususnya terkait dengan aplikasi *e-commerce* seperti Tokopedia dan Shopee. Berikut beberapa aspek teknologi yang dapat menjadi permasalahan:

#### 1. *Sentiment Analysis*

- a. Mengimplementasikan analisis sentimen pada *review* pengguna menggunakan *supervised learning*, yaitu *Naïve Bayes* dan *SVM* [1].
- b. Mengatasi masalah asal memberikan bintang tanpa sesuai dengan *review*nya dan membuat analisis yang akurat terkait dengan sentimen positif atau negatif [3].

#### 2. *Machine Learning*

- a. Memilih model *machine learning* yang sesuai untuk tugas klasifikasi sentimen, dan bagaimana melatih model tersebut dengan data latih yang relevan [1].

- b. Menangani data yang tidak seimbang (imbalance data) yang mungkin terjadi dalam *dataset review* pengguna [3].

### 3. Term Frequency-Inverse Document Frequency (TF-IDF)

- a. Menerapkan metode TF-IDF untuk mengukur frekuensi kata dan informativitas dari sebuah term dalam *review* pengguna [6].
- b. Mengintegrasikan hasil dari TF-IDF ke dalam proses analisis sentimen secara menyeluruh.

### 4. Teknologi Berbasis Web

Flask mampu dalam melakukan pemrosesan *script* Python, kemudian ditampilkan dalam bentuk website[7]. *Script* yang dijalankan yaitu *script* pengolahan data *review* atau melakukan *sentiment analysis* menggunakan algoritma *machine learning* yang telah ditentukan. Tidak hanya itu, tampilan juga dapat divisualisasikan dalam bentuk grafik agar dapat dianalisis dengan lebih mudah.

Permasalahan dalam aspek teknologi ini menjadi dasar untuk merancang solusi yang melibatkan integrasi berbagai teknologi terkini dalam analisis sentimen dan pengelolaan data. Selain itu, pengembangan model yang akurat dan *scalable* serta pengelolaan data yang baik menjadi fokus utama untuk mencapai hasil yang optimal dalam konteks analisis sentimen.

#### 1.1.3 Tujuan Capstone

Adapun tujuan dari pembuatan capstone ini adalah mengatasi kurang efektifnya sistem rating berupa bintang pada aplikasi *Google Play Store* menggunakan algoritma *Naïve Bayes* dan SVM dengan TF-IDF. Sentiment analysis ini nantinya akan memfokuskan pada aplikasi Tokopedia dan Shopee karena merupakan dua aplikasi teratas pada kategori *e-commerce* di *Google Play Store*. Diharapkan dengan adanya komparasi ini, perbedaan sentimen pada masing-masing aplikasi dapat dianalisis dan ditampilkan dengan aplikasi web.

## 1.2 Analisa Solusi yang Ada

Berdasarkan sub-bab sebelumnya, diketahui bahwa metode TF-IDF merupakan metode yang paling sering digunakan untuk ekstraksi fitur dalam analisis sentimen pada *review* aplikasi *e-commerce*. Selain itu, algoritma *Naïve Bayes* dan SVM merupakan algoritma yang paling sering digunakan untuk klasifikasi dalam analisis sentimen pada *review* aplikasi *e-commerce*. Mengenai metrik, akurasi, presisi, *recall*, dan *f1-score* merupakan metrik yang paling sering digunakan untuk evaluasi klasifikasi dalam analisis sentimen pada *review* aplikasi *e-commerce* [6].