

## BAB I PENDAHULUAN

Pada Bab I Pendahuluan terbagi menjadi enam tahap yaitu latar belakang, perumusan masalah, tujuan tugas akhir, batasan tugas akhir, manfaat tugas akhir, dan sistematika penulisan.

### I.1 Latar Belakang

Seiring berjalannya waktu, teknologi informasi pasti terus mengalami perkembangan inovasi dan bertransformasi ke arah yang semakin canggih. Hal itu dibuktikan dengan kehadiran teknologi informasi yang bergerak di bidang perdagangan seperti *electronic commerce (e-commerce)*. *E-commerce* merupakan proses atau kegiatan bisnis transaksi yang melibatkan barang dan jasa menggunakan teknologi informasi (Febriantoro, 2018). *E-commerce* diperkenalkan sekitar 40 tahun yang lalu dalam bentuk yang paling awal. Didirikan oleh mahasiswa Teknik Elektro, Dr. John R. Goltz dan Jeffrey Wilkins, teknologi *compuserver* dibangun dengan memanfaatkan koneksi *dial-up* dan di tahun 1980-an, *compuserver* memperkenalkan beberapa bentuk paling awal dari email dan konektivitas internet kepada publik dan mendominasi lanskap *e-commerce* (Commerce, 2021). Di Indonesia, *e-commerce* diperkenalkan sejak tahun 1996, dengan berdirinya Dyviacom Intrabumi atau D-Net (Jwadm, 2017) sebagai perintis transaksi *online* hingga saat ini.

Di Indonesia sudah banyak perusahaan *e-commerce* yang berkembang, salah satunya adalah Bukalapak. *E-commerce* yang berdiri pada 2010 ini merupakan salah satu *e-commerce* yang diciptakan oleh anak bangsa. Dalam perkembangannya, Bukalapak mampu bersaing dengan *e-commerce* yang didirikan oleh perusahaan asing, salah satunya adalah Lazada. Bukalapak berhasil menduduki urutan ketiga dalam top 10 *e-commerce* yang ada di Indonesia dan masih memiliki peluang untuk terus meningkat dalam aspek pelayanan terhadap *customer*. Data grafik tersebut dikeluarkan oleh Iprice Insight (Iprice, 2021). Bukalapak menduduki urutan ketiga dalam kunjungan *website* setiap bulannya dengan total sebanyak 29.460.000 terhitung dari Quarter 2 (April, Mei, Juni) 2021 seperti yang terlihat pada gambar I-1 di bawah.

Merchant	Monthly Web Visits	AppStore Rank	PlayStore Rank	Twitter
1 Tokopedia	147.790.000	#2	#4	853.000
2 Shopee	126.996.700	#1	#1	6.038.000
3 Bukalapak	29.460.000	#6	#5	215.600
4 Lazada	27.670.000	#3	#2	430.000
5 Bilibili	18.440.000	#8	#7	529.600
6 Bhinneka	6.996.700	#21	#17	67.100
7 Orami	6.260.000	n/a	n/a	5.820
8 Ralali.com Ralali	5.123.300	#26	n/a	2.880
9 JD ID	3.763.300	#7	#6	42.000
10 Zalora	3.366.700	#5	#8	n/a

Gambar I-1 Peringkat *E-Commerce* di Indonesia sumber (Iprice, 2021)

Mendapatkan kepercayaan konsumen dalam aspek keterbukaan, integritas transaksi, dan perlindungan terhadap informasi menjadi salah satu tantangan sekaligus persaingan bagi seluruh *e-commerce* di Indonesia (Bachtiar R. A., 2020). Selain itu dengan bertambahnya jumlah dan nilai transaksi yang dilakukan oleh *e-commerce* tentu mendorong bertambahnya jumlah pemilik usaha pada sektor yang bersangkutan. Hal ini akan menimbulkan persaingan dalam mekanisme pasar yang memacu Bukalapak untuk terus berinovasi menghasilkan produk serta pelayanan yang bervariasi dengan persaingan harga yang menguntungkan bagi pihak produsen maupun konsumen (Hayati, 2021).

Melihat persaingan *e-commerce* yang semakin ketat saat ini membuat Bukalapak ikut bersaing dalam merencanakan dan melakukan berbagai upaya agar dapat bertahan serta meningkatkan kualitas layanan terhadap konsumen. Salah satu caranya adalah dengan melakukan evaluasi dari hasil *review*. Untuk mendapatkan hasil *review* yang digunakan sebagai bahan evaluasi terdapat beberapa teknik atau metode yang dapat dilakukan salah satunya yaitu *text mining*. *Text mining* merupakan penambangan dokumen teks dari website yang berisi komentar, pendapat, ulasan, *feedback*, kritik, dan *review* yang menjadi hal penting, alasan *text mining* menjadi teknik atau metode untuk mendapatkan hasil *review* yang digunakan sebagai bahan evaluasi yaitu karena jumlah dari hasil *review* yang tidak

sedikit serta berasal dari berbagai bentuk sehingga perlu adanya perapian data apabila data dikelola dengan baik maka dapat memberikan keuntungan berupa informasi yang bermanfaat untuk membantu individu atau organisasi dalam pengambilan sebuah keputusan (Budi, 2017). Selain digunakan untuk mengambil sebuah keputusan, melakukan evaluasi dari hasil *review* juga digunakan untuk mengetahui variabel komentar (Luqyana, Cholissodin, & Perdana, 2018). Untuk dapat mengambil sebuah keputusan dari hasil *review* langkah yang perlu diambil salah satunya adalah dengan mengklasifikasi *review* yang bertujuan untuk mengkategorikan data terhadap komentar atau *review* sehingga dapat membantu peneliti atau pelaku usaha dalam menarik kesimpulan terkait kecenderungan komentar (Indriani, 2014). Analisis klasifikasi diaplikasikan untuk mengelompokkan komentar yang bersifat positif, negatif, dan netral dari konsumen sehingga memudahkan tugas perusahaan dalam meninjau kembali kekurangan yang dimiliki (Yosmita, 2018). Dalam proses pengklasifikasian dipisahkan dalam dua kategori yaitu :

- a. Komentar bersifat negatif, bertujuan untuk mengetahui kritik dan saran yang diberikan konsumen untuk dijadikan bahan evaluasi bagi Bukalapak agar dapat memperbaiki pelayanan sehingga mampu menarik minat konsumen.
- b. Komentar bersifat positif, bertujuan untuk mengetahui penilaian konsumen terkait kecenderungan minat jenis produk atau layanan (Mardi, 2016)

Tidak seperti pada komentar positif dan negatif, komentar netral cenderung tidak memiliki kata yang khas sehingga sulit dikenali. Alasan inilah yang membuat komentar yang bersifat netral sulit terdeteksi sebagai kelas netral, karena pada umumnya kalimat bersentimen netral memiliki variasi kata yang lebih beragam dari kelas komentar yang lain (Santoso, Virgini, & Lukito, 2017).

Terdapat banyak metode yang digunakan untuk proses klasifikasi. *Naïve Bayes Classifier* merupakan algoritma yang banyak digunakan untuk proses klasifikasi teks, akan tetapi *Naïve Bayes* hanya memperhatikan probabilitas saja (Claudy, Perdana, & Fauzi, 2018) Karena pada teori ini, satu probabilitas saja tidak bisa mengukur seberapa dsalam tingkat keakuratannya (Pratiwi, Awangga, &

Setyawan, 2020). Berdasarkan penelitian yang berjudul “*Klasifikasi Data Forum Dengan Menggunakan Metode Naive Bayes Classifier*” klasifikasi data forum dilakukan secara otomatis menggunakan algoritma Naive Bayes berdasarkan probabilitas dan teorema bayesian sampai memperoleh hasil akurasi sebesar 73% (Indriani, 2014). Random Forest merupakan algoritma yang menilai hasil *voting* dari *tree* yang terbentuk. Berdasarkan penelitian yang berjudul “*Komparasi Metode KNN, SVM, Dan Random Forest Pada E-commerce Shopee*” merupakan analisis sentimen yang mengklasifikasikan seberapa besar kepuasan pengguna akun Shopee berdasarkan *review* pada Google Playstore dengan menggunakan algoritma salah satunya adalah Random Forest memperoleh hasil akurasi sebesar 83%. K-Nearest Neighbor (KNN) merupakan algoritma yang memperhatikan kemiripan objek dengan kelas. KNN melakukan klasifikasi dengan melihat kemiripan objek dengan kelas (Claudy, Perdana, & Fauzi, 2018). Berdasarkan penelitian yang berjudul “*Penerapan Text Mining pada Aplikasi Tokopedia Menggunakan Algoritma K-Nearest Neighbor*” juga merupakan analisis yang mengklasifikasikan tingkat kepuasan pelanggan Tokopedia menggunakan algoritma KNN dan memperoleh hasil akurasi sebesar 84,05% (Afdal & Elita, 2022).

Berdasarkan deskripsi di atas penelitian ini akan melakukan analisis klasifikasi hasil *review customer* Bukalapak dengan menggunakan algoritma K-Nearest Neighbor (KNN) untuk mendapatkan hasil akurasi yang lebih baik karena melihat jarak antara objek dengan kelas untuk mengetahui kecenderungan kategori komentar guna meningkatkan layanan.

## **I.2 Perumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan sebelumnya, maka berikut adalah rumusan masalah yang akan dibahas dalam penelitian:

- a. Bagaimana proses penerapan analisis klasifikasi pada *platform* Bukalapak terhadap *review* pengguna dengan algoritma K-Nearest Neighbor?
- b. Bagaimana hasil evaluasi dari analisis klasifikasi *review* pengguna pada *platform* Bukalapak?

- c. Bagaimana perbandingan dari hasil klasifikasi komentar positif dan negatif pengguna pada *platform* Bukalapak ?

### **I.3 Tujuan Tugas Akhir**

Berdasarkan perumusan masalah yang telah disebutkan, maka tujuan dari penelitian yang akan dilakukan adalah:

- a. Untuk mengetahui proses penarapan analisis klasifikasi *review* pengguna dengan algoritma K-Nearest Neighbor.
- b. Untuk mengetahui hasil evaluasi dari analisis klasifikasi *review* pengguna pada *platform* Bukalapak.
- c. Untuk mengetahui perbandingan dari hasil klasifikasi komentar positif dan negatif pengguna pada *platform* Bukalapak.

### **I.4 Batasan Tugas Akhir**

Adapun batasan masalah untuk penelitian yang ditentukan untuk mencegah perluasan pembahasan dalam penelitian, sebagai berikut:

- a. Bahasa yang digunakan peneliti untuk komentar/*review* pengguna adalah bahasa Indonesia.
- b. Variabel kategori klasifikasi yang digunakan adalah komentar negatif dan komentar positif.
- c. Data yang diambil dari hanya data yang berbentuk komentar customer berupa teks dan tidak berupa gambar maupun video pada *platform marketplace* Bukalapak.
- d. Data yang digunakan untuk klasifikasi dan evaluasi adalah data yang telah di SMOTE.
- e. Tidak melakukan perbandingan hasil akurasi antara sebelum dan sesudah proses *resampling* (SMOTE).
- f. Evaluasi hanya dilakukan untuk model terbaik dari *training* model.

### **I.5 Manfaat Tugas Akhir**

Manfaat tugas akhir ini sebagai berikut:

a. Bagi Pengguna,

Sebagai referensi untuk mengetahui pandangan pengguna Bukalapak dari komentar akun Bukalapak sebagai bahan pertimbangan sebelum berbelanja di aplikasi Bukalapak.

b. Bagi Pengembang,

Sebagai bahan untuk dijadikan evaluasi dari segi pelayanan (*service*) terhadap *customer* dalam upaya meningkatkan pelayanan dan performa Bukalapak ke arah yang lebih baik di waktu yang akan datang.

c. Bagi Peneliti,

Sebagai bahan referensi untuk mengetahui hasil analisis apabila ingin melanjutkan ke penelitian berikut dengan algoritma K-Nearest Neighbor. Selain itu juga dapat digunakan sebagai bahan referensi untuk mengetahui cara penerapan analisis klasifikasi dengan algoritma K-Nearest Neighbor.

## **I.6 Sistematika Penulisan**

Tugas akhir ini diuraikan dengan sistematika penulisan sebagai berikut:

### **Bab I Pendahuluan**

Pada bab ini berisi penjelasan mengenai latar belakang penelitian, perumusan masalah, tujuan penelitian, batasan masalah penelitian, manfaat penelitian dan sistematika penulisan.

### **Bab II Tinjauan Pustaka**

Bab ini berisi literatur yang relevan dengan permasalahan yang diambil dan dibahas pula hasil-hasil referensi buku/ penelitian/ referensi lainnya yang dapat digunakan untuk merancang dan menyelesaikan masalah dan untuk menentukan metode/kerangka kerja yang akan digunakan di tugas akhir ini.

### **Bab III Metodologi Penyelesaian Masalah**

Pada bab ini menjelaskan langkah-langkah dari penelitian secara rinci meliputi: sistematika penyelesaian masalah, pengumpulan dan mengolah data, dan metode analisis.

#### **Bab IV Perancangan Sistem Terintegrasi**

Pada bab ini, menjelaskan perancangan dan pengolahan data sebelum di proses meliputi pengumpulan data, *preprocessing data*, perhitungan TF-IDF, SMOTE, splitting data menjadi data *training* dan data *testing*.

#### **Bab V Analisa Hasil dan Evaluasi**

Pada bab ini, disajikan secara mendetail mengenai hasil klasifikasi dengan K-Nearest Neighbors, evaluasi performasi, dan visualisasi dari analisis klasifikasi.

#### **Bab VI Kesimpulan dan Saran**

Pada bab ini dijelaskan kesimpulan dari penyelesaian masalah yang dilakukan serta jawaban dari rumusan permasalahan yang ada pada bagian pendahuluan. Saran dari solusi dikemukakan pada bab ini untuk tugas akhir selanjutnya.