

BAB I PENDAHULUAN

Pada bab pendahuluan berisi uraian mengenai latar belakang penelitian, perumusan masalah, tujuan tugas akhir, batasan tugas akhir, manfaat tugas akhir, dan sistematika penulisan. Pada sub bab latar belakang menjelaskan mengenai deskripsi objek penelitian, isu atau problematika yang diangkat dalam penelitian ini, dan analisa permasalahan. Pada sub bab rumusan masalah menunjukkan pertanyaan yang ditanyakan dan akan dijawab di penelitian. Pada sub bab tujuan penelitian menguraikan aktivitas utama yang dilakukan dalam penelitian untuk dapat menjawab pertanyaan penelitian. Sub bab batasan masalah berisi batasan dari penelitian. Sub bab manfaat tugas akhir menjelaskan manfaat yang diperoleh jika permasalahan dapat diselesaikan dan sub bab sistematika penulisan menjelaskan secara singkat isi dari tugas akhir meliputi uraian masing-masing bab

I.1 Latar Belakang

Teknologi saat ini berkembang pesat ke arah digital. Kemajuan teknologi, komputer dan telekomunikasi telah mendukung perkembangan teknologi internet. Menurut survei yang dilakukan oleh APJII (Asosiasi Penyelenggara Jasa Internet Indonesia) jumlah pengguna internet di Indonesia mencapai 196,7 juta dari tahun 2019 hingga kuartal kedua 2020. Jumlah ini setara dengan 73,7% dari jumlah penduduk Indonesia sebanyak 266,91 juta jiwa. Di masa pandemi, transaksi belanja online meningkat pesat. Berdasarkan data Bank Indonesia (BI), jumlah transaksi *e-commerce* meningkat hampir dua kali lipat di masa pandemi Covid-19. Dari 80 juta transaksi di 2019 menjadi 140 juta transaksi di tahun 2020 (Ariyanti, 2020).

Didirikan pada tahun 2010, Bukalapak secara signifikan mentransformasi cara masyarakat Indonesia dalam menjalani aktivitas ekonomi. Bukalapak merupakan *startup e-commerce* kedua di Indonesia yang meraih predikat sebagai *unicorn*. Dilansir dari CBInsights, valuasi Bukalapak bahkan mencapai US\$ 2,5 miliar atau setara Rp 35 triliun (Astutik, 2019). Bukalapak juga menempati urutan ketiga dalam top 10 *e-commerce* Indonesia yang dikeluarkan oleh Iprice Insight (iprice.co.id, 2021). Seperti terlihat pada gambar I.1 di bawah ini.

Merchant	Monthly Web Visits	AppStore Rank	PlayStore Rank	Twitter	Instagram	Facebook	Number of Employees
1 Tokopedia	147,790,000	#2	#4	853,000	3,828,300	6,525,650	4,944
2 Shopee	126,996,700	#1	#1	603,800	7,757,940	21,855,970	12,192
3 Bukalapak	29,460,000	#6	#5	215,600	1,661,140	2,518,990	2,316
4 Lazada	27,670,000	#3	#2	430,000	2,975,370	31,364,410	4,126
5 Blibli	18,440,000	#8	#7	529,600	1,622,480	8,598,260	1,979
6 Bhinneka	6,996,700	#21	#17	67,100	42,280	1,036,230	487
7 Orami	6,260,000	n/a	n/a	5,820	6,040	351,770	211
8 Ralali	5,123,300	#26	n/a	2,880	41,160	91,390	176
9 JD ID	3,763,300	#7	#6	42,000	590,670	939,770	1,185
10 Zalora	3,366,700	#5	#8	n/a	720,090	7,989,080	580

Gambar I.1 Peringkat E-commerce Bukalapak di Indonesia

Dengan adanya pemeringkatan tersebut, pihak Bukalapak dapat meningkatkan kualitas dan kuantitas layanannya dengan cara mengetahui hasil *customer review* terhadap Bukalapak. Sehingga dengan adanya *customer review* tersebut, pihak Bukalapak dapat melakukan evaluasi terhadap layanan yang diberikan kepada *customer* sehingga nantinya layanan yang diberikan kepada customer akan semakin baik dan dapat meningkatkan pemeringkatan.

Permasalahan yang dihadapi adalah *review* dari *customer* Bukalapak yang terlalu banyak sehingga sulit dan membutuhkan waktu yang lama dalam mengklasifikasi dan menganalisis *customer review* jika klasifikasi tersebut dilakukan dengan cara manual. Oleh karena itu dibutuhkan suatu metode yang dapat mengolah data *review* tersebut dengan cara cepat untuk mengklasifikasikan *customer review*. Metode yang digunakan untuk mengklasifikasikan *review* tersebut adalah analisis klasifikasi. *Review* tersebut nantinya akan diklasifikasi menjadi dua jenis yaitu *positive review* dan *negative review* karena dokumen uji bersentimen netral sulit terdeteksi sebagai kelas sentimen netral, karena pada umumnya kalimat bersentimen netral memiliki variasi term yang lebih beragam daripada kelas sentimen yang lain. Tidak seperti pada kelas sentiment positif dan negative, kelas sentiment netral cenderung tidak memiliki term yang khas, sehingga sulit dikenali (Santoso, Virginia, & Lukito, 2017). Sehingga nantinya *review* dari *customer* tersebut dapat dimanfaatkan sebagai pertimbangan pengguna dalam

menggunakan layanan di Bukalapak, ataupun pihak Bukalapak untuk evaluasi peningkatan kualitas dan kuantitas.

Metode klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data. Tujuannya adalah untuk memperkirakan kelas dari suatu objek yang labelnya tidak diketahui (Fadlan, Ningsih, & Windarto, 2018). Terdapat beberapa algoritma dalam proses klasifikasi, diantaranya Support Vector Machine (SVM), Naive Bayes, k-Nearest Neighbor (KNN), Decision Tree, dan Artificial Neural Networks (ANN). Naive Bayes merupakan teknik prediksi berbasis probabilitas sederhana yang berdasarkan pada penerapan teorema Bayes dengan asumsi independensi yang kuat (Fadlan, Ningsih, & Windarto, 2018). Decision Tree menggunakan representasi struktur pohon (*tree*) di mana setiap *node* merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut, dan daun merepresentasikan kelas (Azwanti, 2018). Sedangkan SVM merupakan algoritma *machine learning* yang bekerja atas prinsip *Structural Risk Minimization* (SRM) dengan tujuan menemukan hyperplane terbaik yang memisahkan dua buah class pada input space (Rahutomo, Saputra, & Fidyawan, 2018) dan merupakan metode yang banyak digunakan untuk klasifikasi (Somantri, 2016).

Beberapa penelitian terkait dengan klasifikasi teks telah banyak dilakukan sebelumnya. (Nurajijah & Riana, 2019) telah melakukan penelitian dengan judul penelitian Algoritma Naïve Bayes, Decision Tree, dan SVM untuk Klasifikasi Persetujuan Pembiayaan Nasabah Koperasi Syariah. Pada penelitian ini SVM untuk klasifikasi data pembiayaan nasabah koperasi syariah ke dalam kelas macet dan lancar menghasilkan kinerja terbaik dengan akurasi 89,86% dibandingkan Naïve Bayes dan Decision Tree yaitu Decision Tree sebesar 89,02% dan Naïve Bayes sebesar 77,29%. (Wiyono & Abidin, 2018) juga melakukan penelitian dengan judul perbandingan Algoritma Machine Learning SVM dan Decision Tree untuk Prediksi Keaktifan Mahasiswa Berdasarkan hasil penelitian diperoleh bahwa SVM memberikan hasil akurasi terbaik yaitu 95% sedangkan Decision Tree sebesar 93%. Penelitian lain (Riyanto, 2018) telah melakukan penelitian dengan judul Analisis perbandingan algoritma Naïve Bayes dan Support Vector Machine dalam mengklasifikasikan jumlah pembaca artikel online . Dari kedua

metode yang digunakan dalam penelitian, SVM memiliki hasil akurasi yang terbaik yaitu 63,39% sedangkan Naïve Bayes sebesar 57,10%. (Arifin, 2018) telah melakukan penelitian dengan judul analisis perbandingan tingkat performansi metode SVM dan Naïve Bayes Classifier untuk Klasifikasi Jalur Minat SMA yang membandingkan antara dua buah algoritma, yaitu Naïve Bayes dan SVM. Dari kedua metode yang digunakan dalam penelitian, SVM memiliki hasil akurasi yang terbaik yaitu 97.01% sedangkan Naïve Bayes sebesar 90.68%.

Berdasarkan uraian di atas, penelitian ini akan melakukan penelitian yang berjudul “Klasifikasi Review Customer di E-Commerce Bukalapak Menggunakan Metode Support Vector Machine (SVM)” karena pada penelitian-penelitian sebelumnya, metode ini memiliki nilai akurasi yang cukup tinggi.

I.2 Perumusan Masalah

Berdasarkan uraian pada latar belakang, maka dapat diidentifikasi permasalahan sebagai berikut:

- a. Bagaimana cara penerapan analisis klasifikasi terhadap *customer review* dengan menggunakan metode Support Vector Machine (SVM)?
- b. Bagaimana hasil evaluasi dari analisis klasifikasi terhadap *customer review* Bukalapak?
- c. Bagaimana perbandingan dari hasil klasifikasi review positif dan negatif *customer* bukhalapak?

I.3 Tujuan Tugas Akhir

Berdasarkan rumusan masalah sebelumnya, maka penelitian ini bertujuan untuk:

- a. Untuk mengetahui cara penerapan analisis klasifikasi terhadap *customer review* Bukalapak dengan menggunakan algoritma *Support Vector Machine* (SVM)?
- b. Untuk mengetahui hasil evaluasi dari analisis Klasifikasi terhadap *customer review* Bukalapak
- c. Untuk mengetahui perbandingan dari hasil klasifikasi *review* positif dan negatif *customer* bukhalapak

I.4 Batasan Tugas Akhir

Batasan dalam penelitian ini adalah :

- a. Data set yang digunakan didapatkan dari kaggle
- b. Ulasan yang digunakan adalah ulasan Bahasa Indonesia dan Bahasa Inggris
- c. Menggunakan kernel linier pada algoritma SVM
- d. Komentar pada *review* diklasifikasikan menjadi komentar negative dan positif
- e. Tidak melakukan perbandingan hasil akurasi antara sebelum dan sesudah proses SMOTE
- f. Data yang digunakan untuk klasifikasi dan evaluasi adalah data yang sudah melalui tahap SMOTE
- g. Evaluasi hanya dilakukan untuk model terbaik dari proses *training*

I.5 Manfaat Tugas Akhir

Berikut merupakan manfaat dari penelitian ini :

- a. Bagi *Customer*
Hasil dari analisis klasifikasi nantinya dapat dijadikan pertimbangan oleh *customer* ketika membeli produk di *e-commerce* Bukalapak
- b. Bagi Bukalapak
Hasil dari analisis klasifikasi nantinya dapat dijadikan bahan evaluasi dalam meningkatkan layanan *customer*
- c. Bagi Akademik
Hasil dari analisis klasifikasi dapat digunakan sebagai referensi apabila ingin melanjutkan ke penelitian berikutnya menggunakan metode Support Vector Machine

I.6 Sistematika Penulisan

Tugas akhir ini diuraikan dengan sistematika penulisan sebagai berikut:

Bab I Pendahuluan

Pada bab ini berisi uraian mengenai latar belakang penelitian, perumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian

Bab II Tinjauan Pustaka

Pada bab ini berisi literatur yang relevan dengan permasalahan yang diteliti dan dibahas juga hasil-hasil penelitian terdahulu

Bab III Metodologi Penelitian

Pada bab ini dijelaskan langkah-langkah penelitian secara rinci meliputi: kerangka pemecahan masalah, sistematika penyelesaian masalah, tahap pengumpulan data, tahap pengolahan data, dan metode evaluasi

Bab IV Analisis dan Perancangan

Pada bab ini, menjelaskan analisis studi kasus dan design perangkat lunak yang meliputi pengumpulan data, data preprocessing, perhitungan TF-IDF, klasifikasi SVM, dan evaluasi performansi

Bab V Analisis Hasil dan Evaluasi

Pada bab ini membahas secara terperinci mengenai hasil dari klasifikasi SVM dan hasil dari evaluasi performansi

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.