

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Internet telah menjadi bagian dari kehidupan kita sehari-hari, hal ini ditandai dengan perubahan cara berbisnis dan berinteraksi dengan pembeli, penjual dan pembeli sendiri menikmati kenyamanan dalam berbelanja secara *online* dan bertumbuhnya *electronic commerce (e-commerce)*. Media sosial menjadi platform yang efisien untuk bersosialisasi dan juga bertukar informasi, yang kemudian memunculkan jenis *e-commerce* baru yang dinamakan *social commerce (s-commerce)*. *S-commerce* merupakan bentuk perdagangan elektronik baru dan merupakan pengembangan dari *e-commerce*, yang memanfaatkan media sosial dan teknologi 2.0 yang berfokus lebih ke personal dan interaktif, namun tetap memfasilitasi proses penjualan dan pembelian suatu produk. Pada platform *s-commerce*, pembeli dapat berinteraksi satu sama lain dan dapat mencari informasi yang diperlukan dalam membuat keputusan pembelian. Salah satu platform *s-commerce* yang populer di Indonesia adalah TikTok [1].

Menurut laporan We Are Social per April 2023, TikTok merupakan *platform* media sosial terpopuler di dunia dengan 1,09 miliar pengguna aktif di seluruh dunia. Indonesia berada di peringkat kedua setelah Amerika Serikat dengan jumlah pengguna aktif sebanyak 113 juta [2]. Hal ini dimanfaatkan oleh para produsen, terutama para pebisnis Usaha Mikro Kecil Menengah (UMKM) untuk berjualan secara *online* menggunakan media sosial TikTok sebagai *platform s-commerce* dan juga sebagai media promosi digital untuk menemukan pelanggan baru.

TikTok menyediakan platform ideal bagi pembeli untuk dapat turut berpartisipasi dalam memasarkan suatu produk atau layanan dengan berbagi pengalaman pribadi dan tentang informasi tentang produk atau layanan terkait melalui fitur komentar. Komentar yang diberikan oleh pembeli dapat menjadi informatif dan sangat berharga dalam membantu penjual dalam memahami respons pembeli terhadap

produk yang ditawarkan, fitur komentar juga memungkinkan adanya interaksi antara penjual dan pembeli. Bagi pembeli, komentar di TikTok dapat menjadi sumber informasi yang berguna untuk pertimbangan dalam membeli suatu produk. Namun sayangnya, tidak semua komentar berkaitan tentang produk yang ditawarkan tersebut. Salah satu komentar yang sangat mengganggu adalah komentar *spam*. *Spam* merupakan sebuah pesan yang tidak diinginkan dan tidak berhubungan dengan topik yang ada atau pesan yang sama yang dikirim oleh seorang pengguna lebih dari sekali dalam waktu yang singkat dan mengakibatkan ketidaktepatan informasi yang diperoleh [3].

Pada penelitian ini, peneliti melakukan analisis komentar pada akun UMKM yang menggunakan media sosial TikTok sebagai platform *s-commerce*. Komentar tersebut akan dianalisis dan dibagi menjadi komentar potensial, komentar tidak potensial, dan komentar netral. Komentar potensial merupakan komentar pembeli dan calon pembeli yang menunjukkan ketertarikan dengan produk UMKM yang ditawarkan dan bertujuan untuk melakukan pembelian. Sementara komentar potensial merupakan kebalikannya, yang menunjukkan ketidaktertarikan pembeli dan calon pembeli tentang produk yang ditawarkan. Lalu ada komentar netral, yang menunjukkan ketidaksesuaian dengan produk yang ditawarkan, komentar *spam* akan masuk ke dalam kategori komentar ini. Sikap positif dan negatif dapat membentuk minat seseorang untuk membeli produk. Sikap positif ditunjukkan dari adanya ketertarikan untuk mencari informasi tentang produk yang dijual dan melanjutkan untuk membeli produk tersebut. Sementara sikap negatif ditunjukkan dari tidak adanya ketertarikan untuk mencari informasi tentang produk yang dijual dan tidak berminat untuk membeli produk tersebut [4]. Menurut [5], terdapat hubungan yang positif antara komentar dengan niat pembelian produk. Aspek keakraban, kehadiran sosial, kepercayaan, dan pencarian informasi melalui *s-commerce* juga memiliki pengaruh positif langsung pada niat pembelian.

Untuk mengetahui komentar yang dituliskan merupakan masuk ke dalam kelas potensial, tidak potensial, ataupun netral dapat dilakukan dengan melakukan analisis komentar. Kendala dalam menganalisis komentar pada media sosial terutama TikTok adalah banyaknya penggunaan singkatan dan bahasa yang tidak

baku sehingga sulit dipahami. Oleh karena itu, diperlukan pendekatan *Natural Language Processing* (NLP) untuk memperbaiki bahasa pada komentar agar komentar lebih mudah dipahami oleh sistem. Salah satu pendekatannya adalah menggunakan metode berbasis isi teks dengan menggunakan metode seperti algoritma klasifikasi pada teks komentar seperti algoritma Logistic Regression dengan pembobotan Term Frequency-Inverse Document Frequency (TF-IDF). Metode TF-IDF merupakan cara untuk memberi bobot pada hubungan suatu kata (term) dengan suatu dokumen. Metode ini menggabungkan dua konsep untuk kata dalam dokumen tertentu dan frekuensi terbalik dari dokumen yang mengandung kata tersebut. [5] melakukan penelitian yang menganalisis komentar di aplikasi TikTok untuk mengetahui sentimen pengguna, dengan menggunakan algoritma Logistic Regression, nilai akurasi yang dihasilkan adalah sebesar 84%, yang merupakan 2% lebih akurat dari metode SVM dan 5% lebih akurat dari metode Naïve Bayes. Berdasarkan penelitian [6], algoritma *Term Frequency-Inverse Document Frequency* (TF-IDF) digunakan untuk menganalisis komentar potensial. Algoritma ini berhasil digunakan untuk pengambilan artikel yang relevan pada situs berita online menggunakan kata kunci pencarian, dengan nilai recall 1 dan nilai rata-rata presisi adalah 0,5. Melalui analisis komentar, pelaku UMKM dapat memperoleh respons pembeli terkait produk yang telah ditawarkan dan juga dapat meningkatkan ketertarikan produk tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah untuk proyek akhir ini adalah:

1. Bagaimana cara pelaku UMKM Kota Bandung mengetahui perilaku konsumen yang tertarik melakukan pembelian dan yang tidak tertarik melakukan pembelian berdasarkan komentar pada aplikasi TikTok?
2. Bagaimana cara menganalisis komentar potensial dan komentar tidak potensial yang berpengaruh pada keputusan pembelian konsumen pada aplikasi TikTok?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah disebutkan, tujuan dari proyek akhir ini adalah:

1. Menganalisis komentar pada akun TikTok UMKM Kyla Food (@kylafood) serta akun-akun TikTok lain yang memberikan ulasan berupa video TikTok mengenai produk Kyla Food lalu mengklasifikasikannya ke dalam kelas potensial, tidak potensial, dan netral.
2. Mengetahui perilaku konsumen terhadap produk yang dipromosikan pada akun TikTok pelaku UMKM (@kylafood).

1.4 Batasan Masalah

Beberapa hal yang menjadi Batasan masalah dalam penelitian proyek akhir ini diantaranya adalah:

1. Analisis sentimen hanya dilakukan pada akun TikTok UMKM Kyla Food (@kylafood) serta akun-akun TikTok lain yang memberikan ulasan berupa video TikTok mengenai produk Kyla Food.
2. Dataset pada penelitian ini hanya berbentuk komentar dengan tiga kategori, yaitu produksi produk, penyajian produk, dan *endorse*.
3. Analisis sentimen yang dilakukan hanya menggunakan satu model algoritma saja yaitu model Logistic Regression.
4. Pelabelan teks komentar menggunakan kamus leksikon InSet [27], bukan leksikon khusus untuk UMKM produk makanan.

1.5 Definisi Operasional

1. Analisis Sentimen : Definisi operasional untuk “Analisis Sentimen” adalah salah satu bagian dari studi *text mining* yang bisa disebut studi komputasi untuk melakukan klasifikasi terhadap pendapat, emosi dan sikap seseorang terhadap entitas. Klasifikasi ini akan memberikan pernyataan positif atau negatif pada sebuah kalimat yang berada di dalam dokumen tertentu [7]. Analisis sentimen juga bisa disebut sama dengan *opinion mining* karena

pada metode ini berfokus pada pendapat yang bersifat positif, netral atau pun negatif [8]. Analisis sentimen digunakan dalam berbagai bidang untuk mengetahui opini masyarakat terhadap suatu isu dan membuat solusi yang tepat berdasarkan data yang dikumpulkan.

2. TikTok : Definisi operasional untuk “TikTok” adalah platform seluler yang dikembangkan oleh industri teknologi besar di Tiongkok, yaitu ByteDance. TikTok memungkinkan pengguna untuk membuat video, memiliki fitur menarik, dan basis data yang komprehensif. Semakin populernya TikTok, tidak menutup kemungkinan bisnis bisa mendapat keuntungan dari penggunaan media sosial ini karena TikTok sedang beralih dari platform media hiburan menjadi alat untuk strategi bisnis [9], oleh karena itu TikTok juga dapat digunakan untuk memasarkan produk-produk ke audiens yang lebih luas.
3. Algoritma Logistic Regression : Definisi operasional untuk “Algoritma Logistic Regression” adalah sebuah algoritma jenis regresi analisis yang dapat digunakan untuk menjelaskan hubungan antara kedua variabel berupa kategori, seperti benar atau salah, besar atau kecil, dan sebagainya [10]. Algoritma ini juga cukup kuat untuk klasifikasi biner dan memiliki performa yang tinggi.
4. Usaha Mikro/Kecil/Menengah (UMKM) : Definisi operasional untuk “Usaha Mikro/Kecil/Menengah (UMKM)” adalah unit usaha yang berdiri sendiri yang dilakukan oleh perorangan atau merupakan badan usaha di semua sektor ekonomi dan bukan merupakan anak perusahaan ataupun menjadi cabang dari perusahaan besar baik secara langsung maupun tidak langsung [11].

1.6 Metode Pengerjaan

Metode pengerjaan pada dokumen proyek akhir ini terdiri dari beberapa tahap, yaitu tahap implementasi yang terdiri dari :

1. Pengumpulan Dataset (*Dataset Collecting*)

Tahap ini merupakan tahap awal pada analisis sentimen, yaitu pengumpulan data yang relevan untuk penelitian ini melalui proses data scraping komentar di platform akun TikTok Kyla Food (@kylafood) dan akun-akun lainnya yang memberikan ulasan kepada produk Kyla Food dalam bentuk video TikTok, *scrapping* dilakukan menggunakan perangkat lunak Apify. Tujuan dari tahap ini adalah mengumpulkan data yang akan digunakan pada tahap selanjutnya. Setelah data terkumpul, dilakukan pengelompokan yang bertujuan untuk memberikan kategori pada setiap teks komentar yang di-*scrapping*, seperti kategori 'Produksi Produk', 'Penyajian Produk' ataupun 'Endorse' .

2. *Pra-pemrosesan Data (Data Pre-processing)*

Tahap ini merupakan tahap pembersihan dan normalisasi data sebagai persiapan untuk analisis lebih lanjut. Tahapan-tahapan yang dilakukan meliputi pembersihan data dengan menghapus simbol, emoji, dan angka (*data cleaning*), *case folding*, menghapus huruf duplikat (*remove duplicate letters*), standarisasi bahasa gaul (*slang word standardization*), tokenisasi (*tokenization*), penghapusan *stopwords*, *stemming*, dan mengembalikannya ke bentuk kalimat semula. Tujuannya adalah untuk memastikan bahwa data berada dalam format yang tepat dan siap untuk dianalisis.

3. Pelabelan dengan InSet *Lexicon*

Tahap ini merupakan tahap memberikan label ke setiap teks komentar dalam dataset. Seperti melabeli teks komentar dengan label 'Potensial', 'Tidak Potensial' ataupun 'Netral' yang tujuannya adalah untuk memberikan "jawaban benar" kepada algoritma yang nantinya akan digunakan untuk belajar.

4. Pembobotan TF-IDF

Tahap ini merupakan tahapan untuk mengubah teks menjadi vektor numerik yang dapat diproses oleh algoritma *machine learning*. Frekuensi kata yang terdapat dalam dokumen (*Term Frequency*) dan indikasi untuk seberapa sering atau jarangya sebuah kata muncul dalam dokumen (*Inverse Document Frequency*). Tahapan ini bertujuan untuk mengubah teks menjadi format yang dapat diproses oleh algoritma.

5. Pelatihan algoritma *Logistic Regression*

Algoritma Logistic Regression merupakan metode klasifikasi yang mampu melakukan prediksi probabilitas suatu kejadian. Pada tahap ini, algoritma ini dilatih menggunakan data dan label yang telah diproses dan dibersihkan sebelumnya, yang memiliki tujuan untuk membangun model yang dapat digunakan untuk melakukan prediksi.

6. Evaluasi

Tahap ini merupakan tahap pengukuran kinerja model yang telah dilatih. Hal ini dapat dilakukan dengan metrik seperti *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*. Tujuannya adalah untuk mengukur seberapa efektif kinerja model dan mengetahui apa saja perbaikan yang perlu dilakukan.

7. Pengembangan Aplikasi *Machine Learning*

Setelah ditemukannya hasil analisis sentimen, dibutuhkan sebuah *dashboard* untuk menampilkan hasil analisis sentimen tersebut dan dilengkapi dengan kesimpulan yang dapat memberikan gambaran besarnya. Oleh karena itu, tahap ini merupakan tahap terakhir untuk memvisualisasikan hasil analisis yang telah dilakukan sebelumnya.

1.7 Jadwal Pengerjaan

Pengembangan proyek akhir ini dimulai dari pengumpulan dataset hingga pengembangan aplikasi machine learning (dashboard). Jadwal pengerjaan proyek akhir dapat dilihat pada Tabel 1. 1.

No	Aktivitas	November 2023				Desember 2023				Januari 2024				Februari 2024				Maret 2024				April 2024				Mei 2024				Juni 2024			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
		1	<i>Research for Bandung City's MSMEs</i>																														
2	<i>Dataset Collecting</i>																																

