

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Social Commerce (s-commerce) merupakan pengembangan dari *e-commerce*, kegiatan perdagangan elektronik yang memanfaatkan media sosial dan teknologi web 2.0. Pada *platform s-commerce*, konsumen dapat berinteraksi satu sama lain dan mencari informasi yang diperlukan untuk membantu mereka dalam membuat keputusan pembelian. Menurut laporan *We Are Social*, TikTok merupakan platform media sosial terpopuler di dunia karena memiliki 1,09 miliar pengguna aktif di seluruh dunia per April 2023. Adapun Indonesia merupakan negara dengan jumlah pengguna TikTok terbanyak kedua setelah Amerika Serikat, yaitu sebesar 112,97 juta pengguna. Jumlah tersebut hanya selisih 3,52 juta pengguna dari jumlah pengguna TikTok di Amerika Serikat. Kondisi ini dimanfaatkan oleh para pebisnis Usaha Mikro Kecil Menengah (UMKM) untuk menggunakan TikTok sebagai platform *s-commerce* dan sebagai media promosi digital untuk menemukan pelanggan baru.

Komentar pada aplikasi TikTok dapat menjadi sangat berharga, informatif, dan sangat membantu penjual dalam memahami respons pengguna TikTok terhadap produk yang ditawarkan. Selain itu, melalui komentar pelanggan dan penjual bisa saling berinteraksi hingga melakukan *closing* penjualan. Manfaat komentar di TikTok bagi pembeli adalah untuk mencari informasi yang diungkapkan oleh pengguna lain pada fitur komentar yang dapat menjadi pertimbangan dalam membeli produk tersebut. Adapun salah satu hal yang dapat mengganggu informasi pada komentar di TikTok adalah spam. Spam adalah sebuah pesan yang tidak diinginkan, tidak berhubungan dengan konten yang diunggah atau pesan yang sama dan dikirim oleh seorang pengguna lebih dari satu kali. Hal ini dibuktikan oleh penelitian yang dilakukan oleh Lukito (2017) bahwa dari 17.007 komentar yang ada di akun TikTok artis Indonesia, 10,719 di antaranya adalah spam. Pada penelitian selanjutnya hasil analisis komentar diklasifikasikan ke dalam dua kelas yaitu

komentar potensial dan komentar tidak potensial. Pada penelitian tersebut, spam termasuk ke dalam kelas komentar tidak potensial [1].

Pada penelitian ini, dilakukan analisis komentar pada akun media sosial TikTok UMKM di Desa Lengkong Bandung, yaitu Raja Ngemil (@ownerrajangemil) untuk mengetahui jenis komentar potensial, tidak potensial, dan netral. Komentar potensial yang dimaksud adalah komentar pengguna yang suka dan tertarik dengan produk UMKM yang ditawarkan serta bertujuan untuk melakukan pembelian. Indikator minat beli dapat dijelaskan oleh komponen-komponen sebagai berikut: 1) Tertarik untuk mencari informasi tentang produk 2) Mempertimbangkan untuk membeli 3) Tertarik untuk mencoba 4) Ingin memiliki produk tersebut. Sedangkan, komentar tidak potensial yang dimaksud adalah komentar pengguna yang kecewa dan memiliki pengalaman buruk terhadap produk dan layanan UMKM yang ditawarkan sehingga tidak ada minat beli atau bahkan berhenti melakukan pembelian.

Penelitian Anjaskara pada tahun 2016 mengenai minat beli pada TikTok mengatakan bahwa sikap positif dan negatif dapat membentuk minat seseorang dalam belanja produk. Sikap positif ditunjukkan dari adanya ketertarikan untuk mencari informasi terhadap produk yang dijual dan setelah itu melakukan pembelian produk. Sedangkan, sikap negatif dapat ditunjukkan dari tidak adanya ketertarikan untuk mencari informasi tentang produk yang diiklankan, tidak berminat untuk membeli, dan menunjukkan ketidaksukaan terhadap produk tersebut. Beberapa penelitian lainnya, yaitu penelitian Al Adwan dan Kokash pada tahun 2019 mengatakan bahwa ada hubungan yang positif antara komentar dengan niat pembelian. Menurut mereka, keakraban, kehadiran sosial, kepercayaan, dan pencarian informasi melalui *s-commerce* memiliki pengaruh positif langsung pada niat pembelian [2].

Penelitian yang dilakukan oleh Fauzi, Akbar, dan Asmawan pada tahun 2019 dalam menganalisis komentar di Instagram untuk mengetahui sentimen publik pada sebuah Universitas, dengan metode klasifikasi menggunakan Naive Bayes dan mendapatkan akurasi sebesar 76,5%. Pada penelitian mereka juga, metode

pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF) digunakan untuk menganalisis komentar potensial [3].

Metode TF-IDF adalah 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. Menurut Khusna dan Agustina pada tahun 2019, metode 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 [1].

Melalui analisis komentar, pelaku UMKM dapat memperoleh respons pengguna pada produk yang ditawarkan. Permasalahan dalam menganalisis komentar pada media sosial adalah banyaknya penggunaan singkatan dan bahasa yang tidak baku sehingga sulit dimengerti. Oleh karena itu, pendekatan *Natural Language Processing* (NLP) diperlukan untuk memperbaiki bahasa pada komentar sehingga komentar lebih mudah dipahami oleh sistem.

1.2 Rumusan Masalah

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

1. Bagaimana cara pelaku UMKM Desa Lengkong mengetahui perilaku konsumen yang tertarik melakukan pembelian, puas terhadap produk, dan yang tidak tertarik melakukan pembelian berdasarkan ribuan komentar pada akun TikTok mereka ?
2. Bagaimana cara menganalisis komentar potensial dan komentar tidak potensial yang berpengaruh pada keputusan pembelian konsumen pada produk UMKM Desa Lengkong ?

1.3 Tujuan

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

1. Menganalisis komentar pada akun TikTok UMKM Desa Lengkong dan mengklasifikasikannya kedalam kelas potensial, tidak potensial, dan netral.
2. Memberikan informasi kepada UMKM Desa Lengkong dari hasil analisis sentimen untuk membantu usaha mereka semakin berkembang dan meningkatkan kualitas produknya berdasarkan komentar yang ada pada video-video pada akun TikTok mereka.

1.4 Batasan Masalah

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

1. Analisis sentimen hanya dilakukan pada 1 UMKM di Desa Lengkong, yaitu **Raja Ngemil**.
2. Dataset pada penelitian ini hanya berupa komentar yang berasal dari akun TikTok Raja Ngemil (@ownerrajangemil).
3. Analisis sentimen dilakukan hanya menggunakan satu model algoritma, yaitu Naïve Bayes.
4. Pelabelan otomatis pada komentar hanya menggunakan InSet Lexicon yang berasal dari data twitter.

1.5 Definisi Operasional

1. Analisis Sentimen : Jenis pengolahan bahasa alami yang digunakan untuk melacak perasaan dan emosi masyarakat tentang produk atau topik tertentu. Analisis sentimen juga dikenal sebagai tambang pendapat, di mana sistem mengumpulkan dan meneliti pendapat tentang produk yang dibuat dalam postingan, tweet, komentar, ulasan, dan blog. Analisis sentimen dapat bermanfaat dalam beberapa hal, seperti dalam pemasaran untuk mengetahui apakah iklan atau peluncuran produk baru berhasil,

menentukan versi produk baru atau jasa yang populer, dan bahkan menentukan apakah populasi suka atau tidak suka terhadap fitur atau kebijakan tertentu [4] .

2. TikTok : Sebuah aplikasi jaringan sosial dan platform video musik asal Tiongkok yang diluncurkan pada September tahun 2016. Aplikasi tersebut dipergunakan para penggunanya untuk membuat video musik berdurasi pendek mereka sendiri. Dilihat dari pengguna aktif Tik Tok sebesar 625 juta menjadikan Tik Tok sebagai sarana pemberian informasi yang cepat dan menarik saat ini. Aplikasi tersebut memberikan akses penggunaanya untuk berpartisipasi, berbagi dan membuat konten menarik mereka sendiri [5].
3. Algoritma Naïve Bayes : Salah satu algoritma klasifikasi probabilistik yang berdasar pada Teorema Bayes. Teorema Bayes adalah suatu teorema probabilitas yang menghitung probabilitas kondisional secara terbalik yaitu kemampuan teorema untuk memperbarui keyakinan tentang suatu peristiwa setelah melihat bukti atau informasi baru. Algoritma Naïve Bayes sering digunakan dalam analisis dan klasifikasi teks atau dokumen [6].
4. Usaha Mikro, Kecil, dan Menengah (UMKM) : Kegiatan usaha yang didirikan oleh warga negara dan dapat berupa usaha pribadi atau badan usaha. Usaha mikro, kecil, dan menengah (UMKM) memainkan peran penting dalam meningkatkan perekonomian masyarakat dan berkontribusi pada peningkatan pendapatan masyarakat. UMKM memainkan peran penting dalam pertumbuhan ekonomi negara [7].

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 pertama pada analisis sentimen, yaitu mengumpulkan data melalui proses *scrapping* komentar di media sosial

TikTok Raja Ngemil (@ownerrajangemil) menggunakan perangkat lunak Apify. Tujuan dari tahap ini adalah mengumpulkan data yang akan digunakan pada tahap selanjutnya. Setelah dataset terkumpul, dilakukan pengelompokan dengan memberikan kategori pada setiap video yang telah dilakukan *scrapping*, seperti kategori 'Mukbang', 'Proses Produksi', dan 'Live Promo'.

2. Pra-pemrosesan Data (*Data Preprocessing*)

Pada tahap ini, data dibersihkan dan dinormalisasi untuk mempersiapkan analisis lebih lanjut. Beberapa proses yang dilakukan termasuk pembersihan data dengan menghapus emoji, simbol, dan angka (*data cleaning*), *case folding*, menghapus duplikasi huruf (*duplicate letter removal*), normalisasi *slang* (*slangs*), tokenisasi (*tokenization*), penghilangan *stopwords*, *stemming*, dan mengubah bentuk token ke dalam kalimat. Tujuannya adalah untuk memastikan bahwa data dalam format yang tepat dan siap untuk dianalisis.

3. Klasifikasi Data (*Data Classification*)

Pada tahap ini, setiap teks komentar dalam dataset diberi label untuk diklasifikasikan menjadi komentar "Potensial", "Tidak Potensial", dan "Netral", menggunakan InSet (*Indonesian Sentiment Lexicon*).

4. Pelatihan Algoritma Naïve Bayes

Algoritma Naïve Bayes merupakan model 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 bertujuan untuk membangun model yang dapat digunakan untuk melakukan prediksi.

5. 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.

6. Pengembangan Aplikasi *Machine Learning*

Setelah ditemukan hasil analisis sentimen, diperlukan sebuah dashboard untuk menampilkan hasil analisis sentimen tersebut dilengkapi dengan

kesimpulannya. Oleh karena itu, tahap ini merupakan tahap terakhir untuk memvisualisasikan hasil analisis sentimen pada tahapan 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 Gambar 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	Research for Lengkong Village's MSMEs																																
2	Dataset Collecting																																
3	Data Pre-processing																																
4	Data Classification																																
5	Analysis and Evaluation																																
6	Develop Dashboard for SA Results																																

Gambar 1. 1 Jadwal Pengerjaan Proyek Akhir