

BAB I PENDAHULUAN

1.1 Latar Belakang

Zaman modern menjadikan masyarakat sering menggunakan media sosial sebagai *microblogging*. Masyarakat menggunakannya untuk mengekspresikan diri dan mengemukakan pendapat. Salah satu media sosial yang digunakan yaitu twitter. Informasi yang dapat dibagikan pada twitter dinamakan *tweet*. *Tweet* memiliki batas yang dulunya 140 karakter menjadi 280 karakter (Haqqizar dan Larasyanti, 2019). *Tweet* dapat berupa teks, foto, video, dan sebagainya, melalui *tweet* pengguna dapat berinteraksi satu sama lain. Mereka dapat membagikan informasi, memberikan pendapat, menjelaskan suatu kejadian dan sebagainya.

Gojek merupakan *startup* yang bergerak dibidang *ride hailing*, Perusahaan Gojek sudah berdiri dari tahun 2010 oleh Nadiem Makarim, hampir seluruh kota di Indonesia sudah terjangkau Gojek. Setelah perkembangan yang cukup pesat, Gojek melebarkan sayapnya ke sektor lain seperti *GoRide, GoCar, GoBluebird, GoFood, GoMart, GoMall, GoSend* dan masih banyak lagi, hal ini menjadikan Gojek sebagai salah satu *startup* terbesar di Indonesia. Aplikasi Gojek telah diunduh sebanyak 190 juta lebih, dengan 2 juta lebih *driver* yang telah terdaftar, dan 900 rb lebih pengguna yang telah menjadi Mitra.

Perusahaan Gojek selalu melakukan pengembangan sistem berdasarkan keluhan atau masukan dari pengguna untuk memperbaiki layanan Gojek agar lebih baik. Twitter merupakan salah satu platform sosial media untuk mendapatkan informasi yang dapat digunakan sebagai pedoman evaluasi dan penilaian suatu layanan, dalam hal ini dibutuhkan pendekatan analisis sentimen dengan mengimplementasikan metode *K-Nearest Neighbor* (KNN) dalam mengklasifikasi sentimen positif, negatif, dan netral.

K-Nearest Neighbor (KNN) termasuk dalam algoritma *supervised learning*. Algoritma ini banyak diaplikasikan pada data mining, *statistical pattern recognition, image processing*. Algoritma *K-Nearest Neighbor* (KNN) dalam memetakan suatu objek berdasarkan atribut dan sampel dari data training sehingga sangat efektif apabila menggunakan data training yang besar untuk hasil lebih akurat. Hal ini diharapkan mampu memberikan informasi tentang respon

pengguna Gojek terhadap layanan yang sudah diberikan oleh Gojek yang nantinya dapat digunakan sebagai pedoman dalam penilaian dan evaluasi layanan Gojek.

1.2 Perumusan Masalah

1. Bagaimana cara melakukan analisis sentimen terhadap pengguna Gojek menggunakan metode *K-Nearest Neighbor* (KNN) ?
2. Bagaimana cara mengklasifikasikan sentimen positif, negatif dan netral terhadap pengguna Gojek?
3. Bagaimana hasil pengujian yang diperoleh metode *K-Nearest Neighbor* (KNN) pada analisis sentimen terhadap pengguna Gojek ?
4. Bagaimana kualitas layanan Gojek setelah dilakukan analisis sentimen terhadap pengguna Gojek menggunakan metode *K-Nearest Neighbor* (KNN) ?

1.3 Batasan Masalah

1. Data diambil dengan kata kunci gojek pada platform twitter berbahasa Indonesia.
2. Menggunakan metode *K-Nearest Neighbor* dan tools yang digunakan yaitu visual studio code.
3. Menggunakan bahasa pemrograman python.
4. Mengklasifikasi sentimen positif, negatif, dan netral.

1.4 Tujuan

1. Mengetahui analisis sentimen terhadap pengguna Gojek menggunakan metode *K-Nearest Neighbor* yang dapat digunakan sebagai pedoman penilaian dan evaluasi layanan Gojek.
2. Mengetahui cara mengklasifikasikan sentimen positif, negatif, dan netral terhadap layanan Gojek.
3. Mengetahui kualitas layanan Gojek setelah dilakukan analisis sentimen terhadap pengguna Gojek.
4. Mengetahui hasil pengujian metode *K-Nearest Neighbor* (KNN) pada analisis sentimen terhadap pengguna Gojek.

1.5 Sistematika penulisan

Sistematika penyusunan skripsi ini terdiri dari lima bab, dengan rincian singkat sebagai berikut:

Bab I Pendahuluan

Bab ini menjabarkan apa saja yang melatar belakangi penelitian yang dilakukan, rumusan permasalahan yang ditemukan, batasan permasalahan pada penelitian ini, tujuan dilakukan penelitian ini dilakukan, dan sistematika penulisan.

Bab II Tinjauan Pustaka

Bab ini menguraikan tentang penelitian sebelumnya yang dijadikan sebagai referensi, dan teori-teori pendukung dalam penelitian ini, metode TF-IDF (*Term Frequency Inverse Document Frequency*) dalam pembobotan, pengenalan algoritma *K-Nearest Neighbor* (KNN), bahasa pemrograman yang digunakan, dan perangkat lunak lain yang digunakan.

Bab III Metodologi

Bab ini menjabarkan tentang perlengkapan, bahan yang dipakai dalam penelitian, atau teknik pengambilan data (*crawling*) dengan twitter API, tahapan *preprocessing* menjadi data yang siap diolah, teknik pengujian metode *K-Nearest Neighbor* (KNN).

Bab IV Hasil dan Pembahasan

Bab ini menguraikan tentang implementasi hasil analisis sentimen dan perancangan yang sudah diuraikan pada bab III, serta melakukan pengujian untuk mengetahui ketepatan analisis sentimen yang dirancang, menemukan kelebihan dan kekurangan pada penelitian ini.

Bab V Penutup

Bab ini tentang kesimpulan berdasarkan penelitian ini dengan menjabarkan hasil penelitian yang didapatkan, saran dari peneliti yang ditujukan kepada pembaca ataupun para peneliti yang ingin mengembangkan penelitian ini.

