

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Memasuki era globalisasi saat ini teknologi berkembang begitu pesatnya, perkembangan dalam segala aspek kehidupan menjadi serba digital yang selalu berhubungan dengan teknologi. Pada dasarnya, teknologi diciptakan untuk dapat mempercepat dan mempermudah segala aktivitas manusia juga memberikan banyak kegunaan bagi penggunaannya. Era digitalisasi saat ini mampu menciptakan berbagai macam teknologi yang canggih dan mudah yang membuat hampir tidak ada bidang kehidupan manusia yang terbebas dari dampak penggunaannya, salah satunya di bidang industri yang bahkan saat ini mampu memberikan perubahan yang signifikan terhadap kehidupan manusia (Adyelia Murti, 2020). Sehingga dengan pesatnya perkembangan teknologi di bidang industri memberikan inovasi-inovasi baru salah satunya yaitu *online marketing* atau pemasaran secara *online*.

Pemasaran *online* merupakan strategi untuk mempromosikan produknya melalui jaringan internet dengan perantara sosial media. Di era modern ini, sudah banyak kita temui perusahaan komersial yang memasarkan produknya melalui *platform* sosial media. Bahkan baru ini aplikasi sosial media penyedia video berdurasi pendek atau dikenal dengan TikTok juga menyediakan fitur *e-commerce*. Hal ini disebabkan berkat maraknya penggunaan TikTok, menjadikan TikTok sebagai sarana promosi yang relevan dengan konsep pengiklanan berdasarkan *trend* video menarik yang memikat orang lain terhadap produk yang dipromosikan.

Menurut (Yuniarti et al., 2020) TikTok menjadi aplikasi yang sedang naik daun di awal kemunculan Covid-19 dengan total unduhan lebih dari 425 juta kali. Hal tersebut menunjukkan bahwa TikTok menjadi satu dari sekian banyak hadirnya media sosial yang banyak disukai masyarakat digital dan tengah viral terutama di lingkungan generasi milenial. Apalagi diiringi dengan kemunculan fitur baru yang semakin mengundang banyak peminat masyarakat digital. Fitur baru tersebut merupakan fitur belanja *online* yang merupakan *platform* sosial media berbasis *e-commerce* atau dikenal dengan *TikTok Shop*.

Menurut (Sulistianti & Sugiarta, 2022) fitur TikTok *Shop* tersebut menjadi unggul karena selain berperan sebagai media sosial juga dapat digunakan sebagai media pemasaran produk oleh para pelaku usaha. Konsumen dapat melakukan transaksi secara *online*, memberikan ulasan sebuah produk, dan memberi penilaian terhadap toko hanya dengan satu genggam saja. Dalam penggunaannya pengguna tidak perlu memakai aplikasi lain untuk menggunakan *platform* tersebut, cukup dengan memasang aplikasi TikTok pelanggan dapat dengan mudah berbelanja sambil mengakses *platform* TikTok *Shop* yang membuat pengalaman belanja lebih mudah (Novita et al., 2021). Fitur ini banyak diminati karena selain aman juga karena banyaknya tawaran diskon yang menggiurkan bagi pengguna. Akan tetapi dibalik kelebihanannya pasti akan ada tanggapan lain dari masyarakat mengenai kekurangan fitur belanja dari TikTok *Shop* ini. Kekurangan tersebut bisa terjadi dari kesalahan penjual ataupun kesalahan teknis dari fitur TikTok *Shop* itu sendiri, seperti kesalahan *seller* dalam mengirimkan produk pesanan, terjadinya penipuan saat transaksi jual beli, kemasan rusak, terjadi *error* dalam menggunakan aplikasi, pengiriman yang lambat ataupun keluhan lainnya (Artanti et al., 2018).

Media Sosial selain berguna untuk mendapatkan informasi juga memiliki peran sebagai ajang mengekspresikan diri (Boham, 2020). Tak heran banyak orang yang mengungkapkan perasaan dan tanggapan mereka terhadap sesuatu melalui perantara media sosial. Dari banyaknya situs media sosial, *platform* twitter yang paling banyak digunakan dan menjadi perangkat komunikasi populer dikalangan masyarakat sebagai sarana mengakses informasi, berbagi pengalaman, emosi, hingga keluh kesah yang mereka rasakan. Cuitan komentar tersebut akan berguna bagi perusahaan sebagai evaluasi agar dapat mengetahui kualitas produk yang mereka tawarkan, sehingga agar dapat mengetahui pengelompokan penilaian baik atau buruknya sesuatu dapat dilakukan dengan analisis sentimen.

Analisis sentimen merupakan sebuah metode penambangan opini terhadap data tidak terstruktur yang berguna untuk menganalisis pendapat seseorang terhadap produk, jasa, instansi, individu yang ada pada media sosial. Hasil analisis tersebut kemudian dikelompokkan apakah opini tersebut tergolong ke dalam sentimen positif, sentimen negatif atau sentimen netral (Tri Romadloni et al., 2019).

Hingga saat ini analisis sentimen merupakan topik penelitian yang banyak dibutuhkan dan sedang ramai dilakukan (Saputra, 2019). Analisis sentimen berfokus untuk melakukan klasifikasi opini dari suatu data teks dan merupakan cabang ilmu dari *text mining* atau penambangan teks.

Terdapat beberapa penelitian terkait di bidang *text mining* tentang analisis sentimen yang berhasil mengklasifikasi emosi seseorang. Salah satunya penelitian yang dilakukan oleh (Mulyani et al., 2021) dengan judul “Pengaruh N-Gram terhadap Klasifikasi Buku menggunakan Seleksi Fitur pada *Multinomial Naïve Bayes*” penelitian ini, melakukan pencarian kata pada judul buku dilakukan dengan N-Gram yang kemudian diintegrasikan dengan fitur seleksi *information gain* dengan menguji pengaruh *unigram* dan *bigram* terhadap klasifikasi judul pada algoritma *Multinomial Naïve Bayes*. Hasil pengujian menunjukkan *unigram* memiliki nilai akurasi tertinggi yaitu 74,4%. Selain itu penelitian terdahulu oleh (Hamdana, 2021) dengan judul “Pengembangan Sistem Analisis Sentimen Berbasis Java pada Data Twitter Terhadap *Omnibus Law* Menggunakan Algoritma *Naïve Bayes* dan *K-Nearest Neighbor (K-NN)*”. Dalam penelitiannya melakukan analisis pada data twitter mengenai UU cipta kerja *omnibus law* dengan menerapkan metode *Naïve Bayes* dan KNN yang diterapkan dalam sistem berbasis *java* untuk membandingkan tingkat akurasi dari kedua metode tersebut, dihasilkan bahwa metode *K-Nearest Neighbor (K-NN)* adalah metode yang tingkatan akurasinya lebih tinggi dibandingkan dengan metode *Naïve Bayes* dengan tingkat akurasi sebesar 88%.

Berdasarkan uraian latar belakang di atas, peneliti mencoba melakukan analisa sentimen opini masyarakat di twitter terhadap fitur belanja *TikTok Shop* dengan beberapa teknik tambahan yaitu seleksi fitur dengan menerapkan N-Grams dan *undersampling* untuk mendapatkan data yang seimbang. Dalam penelitian ini menggunakan metode klasifikasi *K-Nearest Neighbor* dan menerapkan pengukuran kesamaan jarak tetangga terdekat dengan *cosine similarity*, yang diharapkan dapat memperoleh hasil identifikasi dengan akurasi yang lebih baik. Sehingga judul untuk penelitian ini adalah “Penerapan N-Gram Terhadap Analisis Sentimen pada Fitur Belanja *TikTokShop* Menggunakan Algoritma *K-Nearest Neighbor*”.

1.2 Rumusan Masalah

Berdasarkan latar belakang, penulis merumuskan beberapa permasalahan dalam penelitian ini sebagai berikut:

1. Bagaimana hasil analisis N-Gram dengan implementasi metode *K-Nearest Neighbor* dalam mengklasifikasikan sentimen data *tweet*?
2. Berapakah nilai akurasi, presisi, dan *recall* yang optimal berdasarkan nilai tetangga k terhadap algoritma *K-Nearest Neighbor*?
3. Bagaimana efektivitas penggunaan teknik *undersampling* dalam menangani kasus *imbalance* data?

1.3 Tujuan

Tujuan yang akan dicapai untuk terlaksananya penelitian ini antara lain:

1. Mengetahui hasil implementasi ketiga fitur N-gram terhadap nilai akurasi yang diperoleh dalam mengklasifikasi sentimen *tweet*.
2. Mengetahui nilai ketetanggaan k yang optimal pada algoritma *K-Nearest Neighbor* berdasarkan nilai akurasi, presisi dan *recall*.
3. Mengetahui performa *undersampling* dalam mengatasi *imbalanced* data.

1.4 Manfaat

Keluaran dari penelitian ini diharapkan dapat bermanfaat bagi semua kalangan, di antaranya:

1. Bagi penulis, mengimplementasikan ilmu-ilmu yang didapatkan di masa perkuliahan khususnya pada mata kuliah *data mining* dan *text mining*.
2. Bagi perusahaan, memberikan pengetahuan mengenai respons masyarakat yang dapat dijadikan evaluasi untuk meningkatkan kualitas pelayanan.
3. Bagi peneliti selanjutnya, hasil dari penelitian diharapkan bermanfaat dan sebagai literatur untuk penelitian selanjutnya khususnya dalam bidang analisis sentimen dengan menggunakan fitur N-Gram.

1.5 Batasan Masalah

Agar penelitian fokus dan berada dalam ruang lingkup masalah penelitian, maka peneliti membatasi dengan beberapa batasan masalah sebagai berikut :

1. Data yang dianalisis berasal dari data komentar twitter.
2. Pencarian data menggunakan kata kunci “tiktokshop”.
3. *Crawling* data dilakukan pada 1 April – 7 Juni 2022 dan diperoleh sebanyak 875 data.
4. Penelitian ini menggunakan bahasa pemrograman *Python* dengan *tools* Jupyter NoteBook untuk menganalisis data.
5. Hasil labelisasi data diklasifikasikan menjadi opini positif dan opini negatif.
6. Skenario uji menggunakan *10-fold cross validation*.
7. Teknik N-Gram yang diimplementasikan yaitu *unigram*, *bigram*, dan *trigram*.
8. Nilai ketetanggaan *K-Nearest Neighbor* yang diterapkan dalam penelitian ini yaitu k3, k-5, k-7 dan k-9.
9. Pengukuran kesamaan tetangga terdekat pada K-NN menggunakan metode *cosine similarity*.
10. Pengukuran kinerja hasil klasifikasi yaitu nilai akurasi, presisi, dan *recall*.