

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Media sosial kini sudah menjadi salah satu pilihan perusahaan untuk menawarkan produk dan jasa yang dimiliki kepada konsumen. Promosi merupakan bagian yang sangat mempengaruhi dalam salah satu usaha dan bisnis, apalagi hal tersebut berhubungan langsung dengan market yang terus bergerak. Pada zaman kemajuan teknologi yang terus berkembang pada saat ini, promosi dilakukan dengan berbagai cara untuk mendapatkan keuntungan tinggi pada usaha atau bisnis yang dimiliki oleh masyarakat. Kemudian promosi juga merupakan salah satu upaya yang digunakan untuk memberitahukan atau menawarkan produk atau jasa kepada calon konsumen untuk melakukan pembelian produk tersebut. Promosi juga bertujuan untuk menciptakan loyalitas merek sebuah produk dengan cara meningkatkan kesadaran pelanggan terhadap produk tersebut.

Kick avenue adalah authentic collectible marketplace online pertama di indonesia yang berdiri sejak tahun 2017. Kick avenue berusaha menciptakan marketplace yang efisien dan terpercaya soal keamanan transaksi jual beli online antara pembeli dan penjual produk. Kick avenue memastikan barang yang dijual 100% original dan autentik seperti yang diharapkan oleh konsumen. Selain itu Kick Avenue juga menyediakan fitur penawaran dan permintaan harga transparan untuk mempertahankan kisaran harga dan produk sesuai pangsa harga pasar yang terkini.

Pada beberapa tahun terakhir Kick Avenue menjadi trend fashion marketplace di Indonesia. Sebelumnya Kick Avenue sendiri hanya merupakan perusahaan baru yang masih dalam proses pengembangan. Namun inovasi yang tiada henti dari tim pengembang dalam mempromosikan sosial media dan didukung dengan ketersediaan produk yang tidak mengecewakan konsumen, sehingga sekarang marketplace ini menjadi salah satu yang dikenali banyak kalangan konsumen di bidang fashion. Diawali dengan kemunculan Kick Avenue, menjadi pintu pembuka bagi berbagai jenis marketplace fashion, sneaker, streetwear lain juga bermunculan. Hingga saat ini, Kick Avenue telah memiliki followers di

platform sosial media sejumlah 533 ribu dan memiliki postingan 3.775 post dalam mempromosikan produk-produk unggulannya[1].

Data mining dapat menjadi salah satu solusi untuk dapat memprediksi perilaku konsumen terutama ketika seseorang mulai memfollow atau like produk atau jasa tertentu. Pengertian lain menyebutkan bahwa data mining adalah sebuah ekstraksi informasi yang tersirat, yang belum diketahui sebelumnya, dan berpotensi bermanfaat dari data. Tidak dapat dipungkiri bahwa setiap hari dalam menjalankan sehari-hari manusia menghasilkan data sehingga semakin lama data tersebut semakin banyak, oleh karena itu penggunaan data mining dapat dimanfaatkan untuk menggali pengetahuan baru dari data yang besar.

Dalam penelitian ini, diharapkan dapat memprediksi kemungkinan konsumen untuk melakukan pembelian dengan melihat karakteristik aktivitas sosial medianya, salah satunya dengan metode K-Nearest. Algoritma K-Nearest Neighbor (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian, berdasarkan klasifikasi data pembelajaran. Sebuah titik pada ruang ini ditandai kelas  $c$  jika kelas  $c$  merupakan klasifikasi yang paling banyak ditemui pada  $k$  buah tetangga terdekat titik tersebut. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan jarak Euclidean.

KNN adalah suatu metode yang menggunakan algoritma supervised dimana hasil dari query instance yang baru diklasifikasikan berdasarkan mayoritas dari label class pada KNN. Tujuan dari algoritma KNN adalah mengklasifikasikan objek baru berdasarkan atribut dan training data. Algoritma KNN bekerja berdasarkan jarak terpendek dari query instance ke training data untuk menentukan KNN-nya.[2]

Berbagai hasil penelitian sebelumnya, telah membuktikan bahwa KNN dapat digunakan sebagai cara untuk melakukan uji dan latih untuk mendeteksi

komentar tertentu. Penelitian berjudul "Implementasi Distance Weighted K-Nearest Neighbor Untuk Klasifikasi Spam & Non-Spam Pada Komentar Instagram" yang dilakukan oleh Antonius Rachmat Chrismanto, Yuan Lukito, Anton Susilo (2020)[3] dimana Metode DWKNN digunakan sebagai perbaikan dari metode KNN melalui pelatihan sistem dengan data latih acak. Setelah proses pelatihan, dilakukan pengujian berdasarkan data uji dan latih dengan parameter nilai k dan persentase fitur yang akan digunakan untuk menguji dan membandingkan metode KNN maupun DWKNN berdasarkan hasil klasifikasinya. Kontribusi penelitian ini menunjukkan bahwa akurasi metode DWKNN lebih baik daripada KNN, perbedaan nilai k ini tidak memiliki dampak yang terlalu berarti dalam klasifikasi komentar spam, dan seleksi fitur (Features Selection) memiliki hasil success rate yang baik pada penggunaan FS antara 80% - 100%. Akurasi optimal dari KNN adalah 82.36% sedangkan menggunakan DWKNN mencapai 91.08% pada FS 80%.

Sementara, berdasarkan penelitian berjudul Analisis Tingkat Penjualan Produk Menggunakan K-Nearest Neighbor (K-Nn) Dan K-Means (Studi Kasus Perusahaan Kayu Elang Perkasa) yang dilakukan Herianto, Nur Syamsiyah, Sapitri Anggraini (2021)[4], Dari penelitian ini akan diuji yang lebih baik apakah menggunakan analisis klasifikasi atau clustering dengan menggunakan data yang sama. Berdasarkan hasil pengujian akurasi kedua metode diperoleh bahwa K-Means memiliki tingkat akurasi yang lebih baik sebesar 78,37% sedangkan K-NN memiliki tingkat akurasi 76,06%.

Sementara, penelitian berjudul Klasterisasi Online Shop Berdasarkan Caption dengan Algoritma Jaro Winkler Distance dan K-Nearest Neighbor yang dilakukan hikmah Jusadi (2021)[5], ditemukan klasterisasi 10 data caption dari tiga kategori yang berbeda. Sebelum proses klasterisasi, dilakukan proses pra pengolahan teks, yaitu case folding, tokenizing, dan stopword removal yang bertujuan untuk mengembalikan bentuk standar dokumen teks, memecah dokumen menjadi kata-kata, dan pembuangan kata umum yang memiliki frekuensi kemunculan yang tinggi. Jaro Winkler Distance digunakan untuk menghitung nilai kesamaan string antara data uji dan data latih dan K-Nearest Neighbor digunakan

untuk proses klusterisasi berdasarkan nilai K yang sudah ditentukan yaitu 3, 53, 103, 153, 203, 249, dan 449. Nilai K dengan hasil akurasi tertinggi adalah K 249. Berdasarkan analisis yang dilakukan dari tujuh kali pengujian dari 450 data latih dan 10 data uji mendapatkan hasil Precision 0,419, Recall 0,567, F-Measure 0,398, dan nilai accuracy 66%.

Sementara, penelitian berjudul "Klasifikasi Komentar Bullying Pada Instagram Menggunakan Metode K-Nearest Neighbor" yang dilakukan Anindya Nanda Rozana (2019)[6], menunjukkan bahwa Metode KNN telah diuji menggunakan 5 nilai k yang berbeda (7, 9, 11, 13, 15) dengan pembagian data 70:30, 80:20 dan 90:10 menggunakan confusion matrix, dengan masing-masing fold yang dibentuk 4, 5 dan 10 sehingga menghasilkan total 95 fold pengujian, pengujian ini menghasilkan akurasi rata-rata sebesar 58,83%. Akurasi tertinggi terdapat pada k 13 dengan perbandingan 90:10 data pada fold ke 6 yaitu sebesar 77%.

Sementara, penelitian berjudul "Komparasi Data Mining Untuk Memprediksi Minat Klien Memilih Produk Asuransi Pendidikan (Studi pada PT.AJB Bumiputera 1912 Karanganyar) yang dilakukan Heria Yunita Putri (2018) [7] menunjukkan bahwa Metode algoritma data mining yang digunakan yaitu metode algoritma Information Gain, Gain Ratio dan Rule Induction. Dalam pengolahan datanya dibantu dengan aplikasi rapidminer untuk membantu mengetahui hasil accuracy, precision, dan recall. Atribut yang digunakan yaitu macam asuransi (X1), cara bayar (X2), uang pertanggungan (X3), premi (X4), dan minat (Y). Hasil dari implementasi penelitian ini memperoleh nilai accuracy, precision, dan recall pada metode information gain dan gain ratio memiliki nilai tinggi yang sama dibandingkan dengan metode rule induction dengan nilai accuracy sebesar 85,69 %, nilai precision sebesar 99,18 %, dan nilai recall sebesar 54,50 %.

Berbeda dengan penelitian-penelitian sebelumnya, Penerapan Data Mining Untuk Memprediksi Minat Beli Pada Media Sosial Instagram @kick Avenue, dengan menguji apakah Jumlah postingan Caption tersebut dapat mengidentifikasi minat beli Ya dan Tidak pada media sosial instagram. Dengan demikian, akan

memudahkan klasifikasi dari KNN dalam menguji niat pembeli.

Dengan demikian, berbagai variasi hasil penelitian mendorong peneliti untuk dapat melakukan analisis mengenai “Penerapan Data Mining Untuk Memprediksi Minat Beli Pada Media Sosial Instagram Dengan Metode K-Nearest Neighbor” sebagai judul dari penelitian ini.

### **1.2 Perumusan masalah**

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

Bagaimana Metode K- Nearest Neighbor dapat digunakan untuk memprediksi minat beli Ya dan Tidak pada media sosial Instagram @kick Avenue?

### **1.3 Tujuan Penelitian**

Tujuan dari penelitian ini adalah untuk memprediksi minat beli pada media sosial Instagram @kick Avenue. Penelitian ini dapat membantu pelanggan dalam memprediksi niat pembelian.

### **1.4 Batasan Masalah**

Adapun Batasan masalah yang digunakan penulis agar pembahasan dalam penelitian ini tidak menyimpang adalah sebagai berikut:

1. Penelitian ini bertujuan untuk mengidentifikasi minat beli Ya dan Tidak dari Jumlah Post Caption Instagram
2. Data yang digunakan adalah Jumlah Post Caption dari akun Instagram official @kick Avenue.
3. Data set yang digunakan untuk penelitian adalah data set primer.

### **1.5 Manfaat Penelitian**

Manfaat Secara akademis, penelitian ini diharapkan akan bisa memberikan manfaat dan kontribusi akademis, diantaranya adalah sebagaimana berikut ini:

1. Hasil penelitian diharapkan menjadi bahan rujukan bagi peneliti lain yang berminat pada kajian mengenai minat beli dengan KNN.
2. Menambah pengetahuan bagi upaya perkembangan ilmu pengetahuan,

khususnya tentang persepsi dalam kajian ilmu IT dan penerapan algoritma.

Secara Praktis, diharapkan penelitian ini dapat memberikan manfaat sehingga dapat mempermudah pelaku bisnis di dalam menyusun strategi pemasaran melalui media sosial, diantaranya adalah:

Secara praktis penelitian ini diharapkan mampu menjadi masukan bagi pihak-pihak yang berkepentingan mengenai penerapan KNN untuk mengetahui sejenis minat beli dan sebagai referensi penelitian lanjutan mengenai permasalahan

#### **1.6 Metode Penelitian**

Metode penelitian yang akan digunakan dalam pembuatan skripsi ini meliputi:

##### **1.6.1 Metode Pengumpulan Data**

Tahap awal dimulai dengan pengumpulan dataset melalui plugin download Instagram pada Google Chrome seluruh data yang ada di akun Instagram @Kick Avenue.

##### **1.6.2 Metode Analisis**

Metode analisis pada penelitian ini menggunakan metode *K Nearest Neighbor* dengan tool berupa *Google Collab* dengan Bahasa pemrograman *python*.

##### **1.6.3 Metode Perancangan**

Sistem dirancang menggunakan bahasa pemrograman *python* dan metode yang dipilih adalah KNN.

##### **1.6.4 Metode Implementasi**

Implementasi metode KNN menggunakan *Google Colab* yaitu platform *jupyter notebook python* pada server *Google*.

##### **1.6.5 Metode Pengujian**

Pada tahap ini dilakukan pengujian prediksi minat, dan di evaluasi hasilnya.