

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Rokok adalah salah satu zat adiktif yang sangat riskan untuk individu dan umum apabila digunakan. Rokok ialah barang yang bersumber dari olahan tembakau yang digulung ataupun dilinting memakai kertas, daun, atau kulit jagung yang mempunyai ukuran selebar kelingking dengan panjang 8-10 cm, lazimnya dihisap bagi penggunaanya selepas dibakar pada ujungnya.

Jumlah perokok dewasa di Indonesia telah meningkat dalam sepuluh tahun terakhir. Menurut hasil Survei Tembakau Dewasa Global (GATS) 2021 yang dirilis oleh Kementerian Kesehatan (Kemenkes), jumlah perokok dewasa meningkat sebanyak 8,8 juta orang, dari 60,3 juta pada tahun 2011 menjadi 69,1 juta pada tahun 2021. Meskipun persentase perokok di Indonesia mengalami penurunan dari 1,8% menjadi 1,6%. Namun, jika kebiasaan merokok setiap tahunnya tidak bisa dikurangi, maka jumlah kematian akibat merokok di Indonesia juga akan terus bertambah.

Beberapa bahan utama yang terdapat dalam rokok meliputi nikotin, benzovrin, metal-klorida, aseton, amonia, dan karbon monoksida. Satu batang rokok mengandung sekitar 4.000 senyawa kimia yang berpotensi berbahaya dan beracun bagi tubuh, di mana 43 di antaranya diketahui memiliki sifat karsinogenik. Nikotin, sebagai salah satu komponen utama, merupakan zat berbahaya yang dapat menyebabkan ketergantungan dan juga memiliki sifat karsinogenik, sementara karbon monoksida dapat menyebabkan penurunan kadar oksigen dalam darah (Sinaga, 2019).

Terdapat 5 variabel yang penulis ambil pada penelitian ini, yaitu jumlah, waktu, tahapan, alasan, usaha berhenti. Variabel jumlah dan waktu yang bersumber dari *website* memiliki 4 kategori yaitu, rendah, rendah-menengah, menengah, dan tinggi. Variabel tahapan yang bersumber dari jurnal memiliki 4 kategori yaitu, tahap *prepatory*, tahap *initiation*, tahap *becoming a smoker*, dan tahap *maintenance of*

*smoking*. Variabel alasan yang bersumber dari jurnal memiliki 4 kategori yaitu, menemani kesendirian, pergaulan, meningkatkan rasa percaya diri, dan pelarian dari masalah. Variabel usaha berhenti memiliki 2 kategori yaitu, ada dan tidak ada.

Fokus dari penelitian ini merupakan mahasiswa dan mahasiswi yang berkuliah di Fakultas Teknik Universitas Malikussaleh. Banyak dari mahasiswa dan mahasiswi Fakultas Teknik Universitas Malikussaleh tidak mengetahui tingkat kecanduan mereka terhadap rokok. Untuk membantu mahasiswa dan mahasiswi Fakultas Teknik Universitas Malikussaleh mengetahui di tingkat mana kecanduan terhadap rokok, maka dibangun sebuah sistem menggunakan salah satu klasifikasi *data mining* menurut cara kerjanya, yaitu *clustering*.

*Data mining* ialah serangkaian sistem yang bertujuan guna mengeksplorasi dan menggali nilai-nilai berupa informasi serta hubungan-hubungan kompleks yang tersimpan pada suatu kumpulan data. Dengan melaksanakan analisis pola informasi terhadap data, tujuannya adalah untuk mengolah data menjelma informasi baru yang lebih berguna melalui proses ekstraksi dan penemuan pola-pola berharga atau menarik yang terdapat pada basis data (Utomo & Purba, 2019).

*Clustering* adalah cara pengelompokan sejumlah objek data menjadi beberapa kelompok atau klaster sehingga objek-objek pada satu kelompok mempunyai kesamaan yang tinggi, tetapi berbeda secara signifikan dengan objek-objek pada kelompok lainnya (Seimahaira, 2021).

Metode yang diaplikasikan pada penelitian *clustering* tingkat kecanduan rokok yang akan dilakukan adalah metode *K-Means*. *K-Means* ialah salah satu cara pengelompokan data di mana setiap titik data pada sebuah klaster ditetapkan oleh tingkat keanggotaannya. Metode ini pertama kali dikembangkan oleh Jim Bezdek pada tahun 1981. *K-Means* ialah algoritma *clustering* yang bersifat iteratif. Algoritma ini dimulai dengan menentukan secara acak jumlah klaster yang akan dibentuk, yang disimbolkan dengan K. Setelah itu, nilai K ditetapkan secara acak, dan titik-titik tersebut dianggap sebagai pusat klaster atau *centroid* menggunakan rumus tertentu hingga jarak terpendek antara setiap data dengan *centroid* terpenuhi. Data kemudian diklasifikasikan berdasarkan kedekatannya dengan *centroid*. Proses

ini diulang hingga tidak ada perubahan pada nilai-nilai *centroid* (D. D. C. Nugraha et al., 2014).

Lokasi penelitian ini akan dilaksanakan di kawasan Universitas Malikussaleh, subjek yang akan menjadi fokus pada penelitian ini ialah mahasiswa dari Universitas Malikussaleh yang diyakini dapat membantu penulis dalam mengumpulkan data yang diinginkan.

Penelitian sebelumnya dilakukan oleh Daniel Andre Marpaung, Murni Marbun dengan judul “Sistem Pendukung Keputusan Penentuan Tingkat Kecanduan Masyarakat Terhadap Rokok dengan Metode Fuzzy Mamdani”. Kriteria yang dipakai untuk penelitian ini yaitu : Biaya, Frekuensi merokok, dan Lingkungan. Hasil yang terdapat dari penelitian ini ialah tingkat kecanduan masyarakat tercatat ada pada angka 13 atau digolongkan sebagai candu, dan Pengaplikasian metode *fuzzy mamdani* pada penentuan tingkat kecanduan masyarakat terhadap rokok mengeluarkan hasil yang akurat sesuai jawaban dari kuesioner yang di isi oleh masing-masing responden (Marpaung & Marbun, 2021).

Penelitian sebelumnya juga dilakukan oleh Ria Tri Wahyuni, A M H Pardede, dan Tioria Pasaribu dengan judul “*Data Mining* Pengelompokan Pecandu Narkoba Di Masyarakat Berdasarkan Faktor Penyebab Penggunaanya Menggunakan Metode *K-Means Clustering*”. Kriteria yang terdapat pada penelitian ini yaitu : Usia, Jenis Narkoba, dan Faktor Penyebab. Hasil dari penelitian Data Mining tentang Pengelompokan Pecandu Narkoba di Masyarakat menggunakan Metode *K-Means Clustering* menunjukkan bahwa terdapat tiga kelompok data pecandu narkoba. Kelompok pertama terdiri dari individu dalam rentang usia 31-37 tahun yang menggunakan Tembakau Gorilla sebagai jenis narkoba, dengan faktor penyebabnya adalah lingkungan. Kelompok kedua terdiri dari individu dalam rentang usia 17-23 tahun yang menggunakan Inex sebagai jenis narkoba, dan faktor pemicunya adalah kepribadian. Sementara itu, kelompok ketiga terdiri dari individu dalam rentang usia 17-23 tahun yang menggunakan lem sebagai jenis narkoba, dengan faktor penyebabnya adalah keluarga.

Penelitian sebelumnya juga dilakukan oleh Wahyu Fuadi, Ar Razi, Dedi Fariadi dengan judul “Automasi Penentuan Tren Topik Skripsi Menggunakan

Algoritma *K-Means Clustering*”. Variabel yang terdapat pada penelitian ini yakni : *Data Mining*, Kecerdasan Buatan, Pengolahan Citra, Sistem Pengambilan Keputusan, dan Sistem Informasi Geografis. Hasil yang didapatkan dari penelitian ini adalah tren topik yang paling populer ialah *Data Mining* sebanyak 48 kategori skripsi, selanjutnya Pengolahan Citra dengan total 43 kategori. Sistem Pengambilan Keputusan menduduki urutan ketiga dengan total 32 kategori. Kecerdasan Buatan dan Sistem Informasi Geografis berada pada urutan ke empat dan lima dengan total 20 dan 14 kategori. Jadi bisa di ambil kesimpulan bahwa tren kategori skripsi yang paling banyak diambil oleh mahasiswa Prodi Teknik Informatika Unimal adalah *Data Mining*. (2022)

Berdasarkan latar belakang tersebut, maka dapat diambil sebuah judul “*Clustering* Tingkat Kecanduan Rokok Pada Mahasiswa Universitas Malikussaleh Menggunakan Metode *K-Means*”.

## **1.2 Rumusan Masalah**

Berlandaskan latar belakang yang telah dijabarkan berikut, maka rumusan masalah yang akan dipecahkan melalui penelitian ini ialah:

1. Bagaimana cara menciptakan sebuah program aplikasi *clustering* tingkat kecanduan rokok memakai metode *K-Means*?
2. Bagaimana mengaplikasikan metode *K-Means clustering* untuk menentukan tingkat kecanduan rokok pada mahasiswa Fakultas Teknik Universitas Malikussaleh?

## **1.3 Batasan Masalah**

Agar pembahasan dan masalah pada penelitian ini kian jelas dan terstruktur, maka masalah penelitian ini akan ditentukan dalam hal-hal berikut:

1. Sistem hanya digunakan untuk menentukan tingkat kecanduan rokok pada mahasiswa Fakultas Teknik Universitas Malikussaleh.
2. Penelitian akan dilaksanakan di kawasan Fakultas Teknik Universitas Malikussaleh.

3. Data yang dipakai merupakan kuisioner yang dibagikan dan diisi oleh mahasiswa Fakultas Teknik Universitas Malikussaleh yang merupakan perokok aktif.
4. Kriteria yang digunakan untuk melakukan *clustering* tingkat kecanduan rokok pada mahasiswa Fakultas Teknik Universitas Malikussaleh adalah jumlah, waktu, tahapan merokok, alasan merokok, dan niatan berhenti merokok.
5. Hasil *cluster* terbagi menjadi 4, yaitu ringan, sedang, berat, dan sangat berat.
6. Sistem ini dibangun menggunakan bahasa pemrograman Python.
7. Sistem memakai algoritma *K-Means* bagi proses *clustering*.

#### **1.4 Tujuan Penelitian**

Berikut menggambarkan tujuan penelitian yang hendak dilaksanakan:

1. Membentuk sistem *clustering* tingkat kecanduan rokok memakai metode *K-Means*.
2. Mengetahui jalan penerapan metode *K-Means* untuk *clustering* tingkat kecanduan rokok.

#### **1.5 Manfaat Penelitian**

Berikut adalah kegunaan yang diinginkan untuk penelitian yang akan dilaksanakan:

1. Setelah terbangunnya sistem *clustering* tingkat kecanduan rokok tersebut, diharapkan dapat membantu untuk mengetahui berapa banyak mahasiswa Universitas Malikussaleh yang tergolong ke dalam perokok tingkat ringan, sedang, berat, atau sangat berat.
2. Untuk mengetahui efektifitas penerapan metode *K-Means* dalam sistem yang akan dirancang.