

## BAB 3

### METODOLOGI PENELITIAN

#### 3.1 Waktu dan Tempat Penelitian

Peneliti memperoleh data dari kuesioner. Dimana penelitian ini dilakukan pada mahasiswa FST UINSU Medan. Penelitian ini dilakukan pada bulan September 2024 sampai selesai.

#### 3.2 Jenis Penelitian

Jenis penelitian yang digunakan adalah kuantitatif. Penelitian kuantitatif menggunakan banyak angka dengan mengumpulkan data untuk kemudian diukur menggunakan metode statistik. Penelitian kuantitatif cenderung melihat fakta berdasarkan data yang dikumpulkan.

#### 3.3 Jenis Data dan Sumber

Jenis data yang digunakan dalam penelitian ini menggunakan kuesioner campuran yang digunakan untuk memperoleh serangkaian data penelitian termasuk yang berbentuk angka-angka. Data yang didapatkan melalui pengisian kuesioner mahasiswa FST UINSU Medan.

#### 3.4 Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan adalah menggunakan kuesioner yang dilakukan dengan cara mengisi pertanyaan pada kuesioner dan membagikan link kuesioner kepada grup mahasiswa saintek tentang kecanduan rokok di FST UINSU Medan. Teori dasar yang digunakan ialah Metode Algoritma *K-Prototypes* agar mengetahui pengelompokan data merokok pada mahasiswa FST UINSU Medan menggunakan metode *K-Prototypes* pada tahun 2024.

#### 3.5 Variabel Penelitian

Variabel penelitian dalam penelitian ini menggunakan data kuesioner mahasiswa FST UINSU Medan pada tahun 2024. Variabel yang menunjukkan data numerik yaitu usia awal merokok dan jumlah rokok (bungkus). Sedangkan variabel yang menunjukkan data kategorik yaitu prodi, jenis rokok, dan faktor merokok.

Tabel 3.1 : Variabel Penelitian

Variabel	Jenis Variabel
$x_1$ : Usia Awal Merokok	Numerik
$x_2$ : Jumlah Rokok (Bungkus)	Numerik
$x_3$ : Prodi	Kategorik
$x_4$ : Jenis Rokok	Kategorik
$x_5$ : Faktor Merokok	Kategorik

### 3.6 Langkah-langkah Penelitian

1. Tentukan jumlah *cluster* ( $k$ ) yang akan dibentuk.
2. Pastikan *centroid* atau titik pusat awal kluster awal menjadi *k prototype* awal.
3. Gunakan persamaan 3.1 untuk menentukan jarak menggunakan ukuran jarak tipe data campuran.

$$d(x_i, c_y) = \sum_{p=1}^m (x_{ip} - c_{yp})^2 + \gamma \sum_{q=1}^m \delta(x_{iq}; c_{yq}) \quad (3.1)$$

Sebelum melakukan persamaan 3.1 tersebut menghitung gamma ( $\gamma$ )

$$\gamma = \frac{1}{m} \sum_{p=1}^m \sigma_p \quad (3.2)$$

4. Semua item dikelompokkan berdasarkan nilai jarak terdekat.

$$e_i = \left\{ \begin{array}{l} 1, s = \min \{d(x_i, c_1), d(x_i, c_2), \dots, d(x_i, c_y)\} \\ 0, \text{lainnya} \end{array} \right\} \quad (3.3)$$

5. Dengan memanfaatkan nilai rata-rata untuk variabel numerik seperti ditunjukkan dalam persamaan 3.4, hitung nilai *centroid* baru.

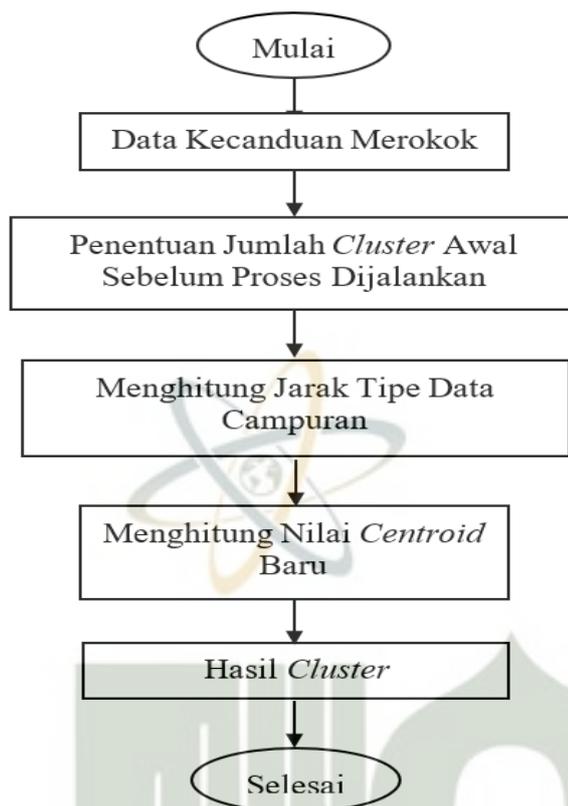
$$c_{yq} = \frac{\sum_{i=1}^{n_y} x_{ip}}{n_y} \quad (3.4)$$

Dengan menggunakan nilai modus, seperti ditunjukkan dalam persamaan 3.5, nilai *centroid* baru untuk variabel kategorik dihitung.

$$c_{yq} = \text{modus} \{x_{1q}, x_{2q}, \dots, x_{iq}\} \quad (3.5)$$

6. Tentukan objek mana yang paling dekat dengan nilai *centroid* baru dan tetapkan kembali ke setiap kluster. Proses algoritma berakhir jika titik pusat *centroid* atau kluster telah konvergen atau tidak lagi berubah.

### 3.7 Prosedur Penelitian



Gambar 3.1 Alur Kerja Prosedur Penelitian