

BAB III

METODE PENELITIAN

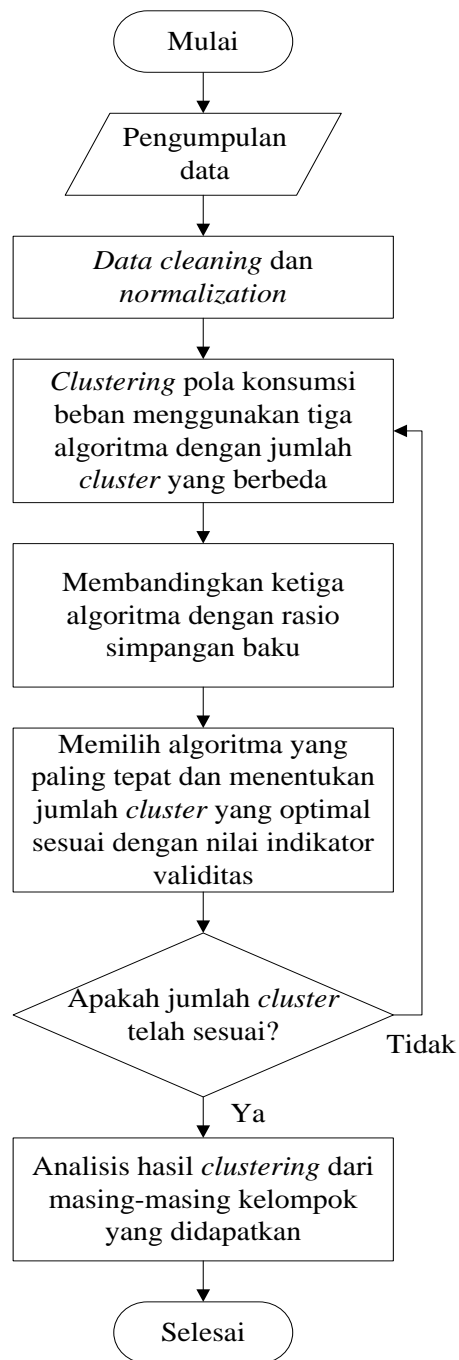
Penelitian ini bertujuan untuk melakukan analisis profil beban listrik selama satu tahun dengan membandingkan tiga buah metode teknik *clustering*. Metode tersebut adalah K-means, Fuzzy c-means, dan K-harmonic means. Untuk menentukan jumlah *cluster* yang optimal digunakan *Davies-Bouldin Index*. Sehingga pada akhirnya dapat mengidentifikasi kebutuhan beban puncak normal dan abnormal pada *cluster* yang terdiri dari beberapa profil beban serupa.

3.1 Sumber Data dan Perangkat Penunjang Penelitian

Data yang digunakan penulis merupakan profil beban harian dengan interval waktu 30 menit selama tahun 2014 yang didapatkan dari P3B Jawa Bali, Region 2 APB Cigereleng. Penelitian yang dilakukan membutuhkan beberapa peralatan untuk menunjang semua proses yang berkaitan dengan teknik *clustering*. Oleh karena itu, digunakan perangkat keras berupa *PC* atau *laptop* dengan spesifikasi *Operating System Windows 7 Ultimate 64-bit (6.1, Build 7601); Processor Intel(R) Core(TM) i3-2328M CPU @ 2.20GHz (4CPUs), ~2.2GHz; Memory 2048MB RAM*. Sedangkan perangkat lunak yang digunakan adalah *Matlab R2009a, Microsoft Office Excel 2016, Mendeley Desktop ver. 1.13.8.0, dan Microsoft Office Visio 2007*.

3.2 Prosedur Penelitian

Langkah-langkah dalam penelitian ini ditunjukkan pada gambar 3.1. Hal pertama yang dilakukan adalah mengumpulkan data profil beban harian selama tahun 2014. Setelah data diperoleh, proses pengelompokan profil beban dilakukan dengan menggunakan ketiga buah metode *clustering* dan mengelompokkannya ke dalam beberapa kelas, yaitu dua hingga sepuluh kelas.



Gambar 3.1 *Flow chart* prosedur penelitian

Setelah proses pengelompokkan, selanjutnya dilakukan evaluasi hasil *clustering* terhadap ketiga metode tersebut. Evaluasi dilakukan dengan membandingkan simpangan baku dalam kelompok dan simpangan baku antar

kelompoknya. Metode yang paling baik merupakan metode yang memiliki rasio simpangan baku paling kecil (Bunkers & Miller, 1995). Simpangan baku dalam kelompok dapat dihitung dengan persamaan (3.1).

$$S_w = \frac{1}{c} \sum_{i=1}^c S_i \quad (3.1)$$

Demikian pula simpangan baku antar kelompok dapat dihitung menggunakan persamaan (3.2).

$$S_b = \left[(c - 1)^{-1} \sum_{i=1}^c (\bar{x}_i - \bar{x})^2 \right]^{\frac{1}{2}} \quad (3.2)$$

Dimana:

\bar{x}_i : rata-rata nilai pada kelompok ke- i

\bar{x} : adalah rata-rata total kelompok.

Tahap selanjutnya menentukan jumlah *cluster* yang optimal berdasarkan indikator validitas *clustering* yang digunakan, yaitu *davies-bouldin index* yang dapat dinyatakan dalam persamaan (3.3) sebagai berikut (Davies & Bouldin, 1979):

$$DBI = \frac{1}{K} \sum_{i=1}^K \max_{i \neq j} (R_{i,j}) \quad (3.3)$$

$R_{i,j}$ adalah ukuran rasio seberapa baik nilai perbandingan antara *cluster* ke- i dan *cluster* ke- j . Nilainya didapatkan dari komponen kohesi dan separasi yang ada pada persamaan (2.7) dan (2.8). *Cluster* yang baik adalah yang mempunyai kohesi yang sekecil mungkin dan separasi yang sebesar mungkin. $R_{i,j}$ diformulasikan oleh persamaan (3.4).

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (3.4)$$

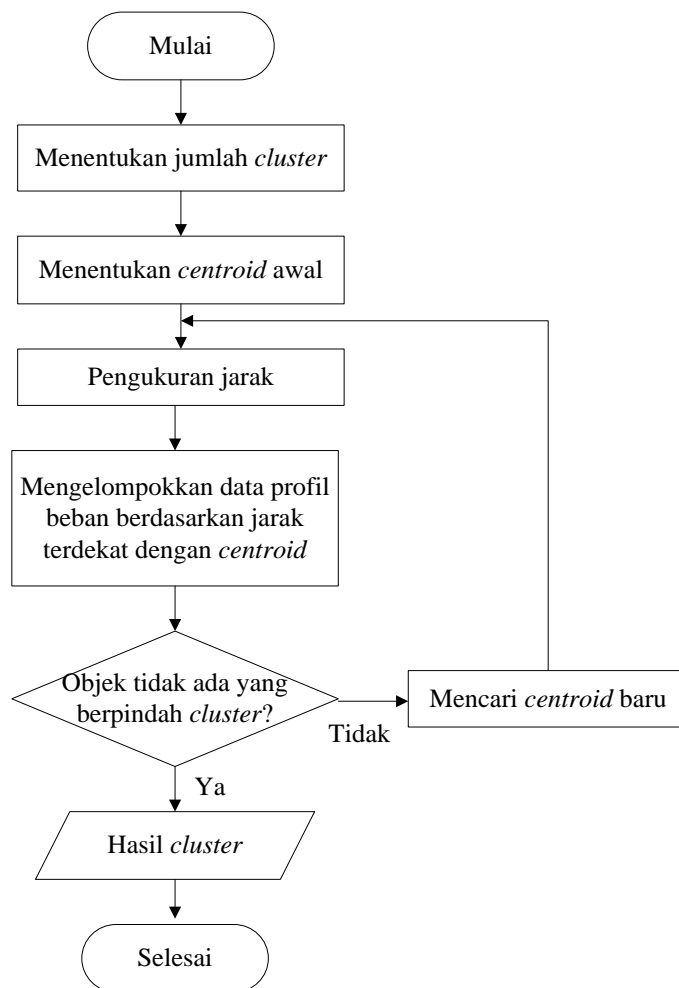
Kriteria banyaknya kelompok yang optimum diberikan oleh nilai DBI yang paling kecil (non – negatif ≥ 0).

Setelah diketahui jumlah kelompok yang optimum, analisis profil beban dilakukan pada masing-masing kelompok dengan mengamati beban puncaknya, sehingga diketahui *load* dan *loss factor* dari masing-masing kelompok. Hingga akhirnya, dari penelitian ini dapat ditarik kesimpulan dan memberikan saran, baik untuk penelitian selanjutnya maupun aplikasi dari hasil yang didapat dalam penelitian.

3.3 Teknik Clustering

3.3.1 Algoritma K-means

Algoritma *K-means clustering* terdiri dari beberapa langkah, seperti yang ditunjukkan pada gambar 3.2.



Gambar 3.2 Flow chart *K-means clustering*

Menentukan jumlah *cluster* dan nilai centroid merupakan hal pertama yang dilakukan dalam algoritma ini (Sharma & Singh, 2014). Dalam menentukan nilai centroid pada awal iterasi dilakukan secara acak. Sedangkan untuk menentukan nilai centroid yang merupakan tahap dari iterasi digunakan persamaan (3.5) sebagai berikut:

$$\bar{v}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj} \quad (3.5)$$

dimana :

- \bar{v}_{ij} : centroid/ rata-rata *cluster* ke-I untuk variable ke-j
- N_i : jumlah data yang menjadi anggota *cluster* ke-i
- i,k : indeks dari *cluster*
- j : indeks dari variabel
- x_{kj} : nilai data ke-k yang ada di dalam *cluster* tersebut untuk variable ke-j

Kemudian menghitung jarak antara titik centroid dengan titik tiap objek. Untuk menghitung jarak tersebut digunakan Euclidean Distance, seperti pada persamaan (3.6).

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (3.6)$$

dimana:

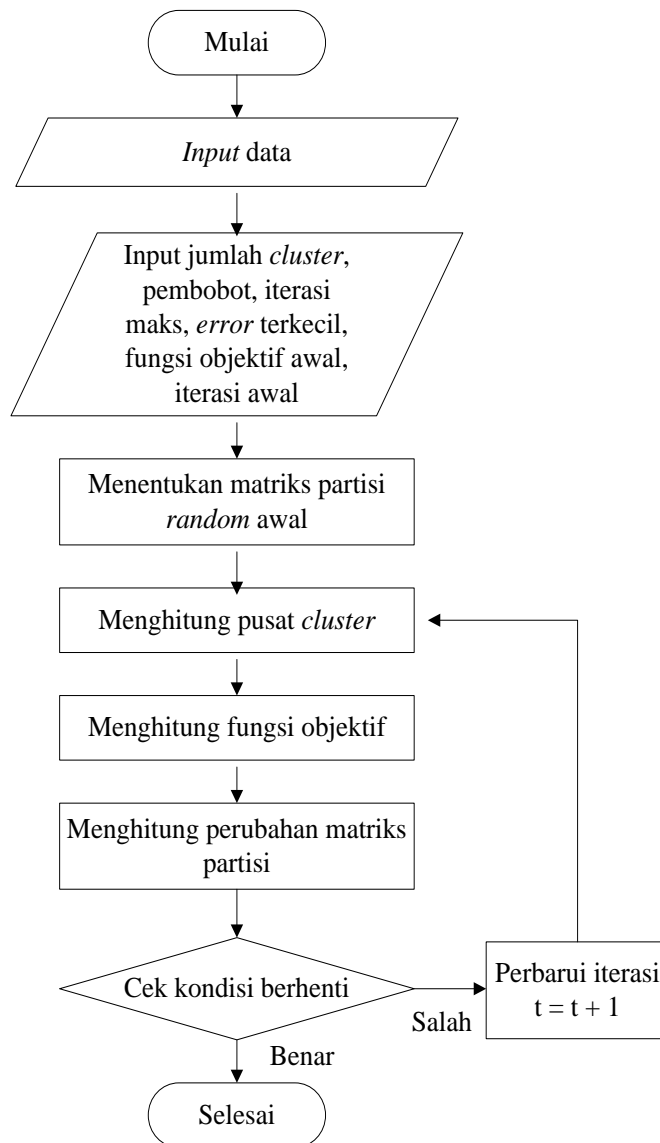
- D_e : Euclidean Distance
- i : banyaknya objek,
- (x,y) : koordinat object dan
- (s,t) : koordinat centroid.

Selanjutnya dilakukan pengelompokkan objek untuk menentukan anggota *cluster* dengan memperhitungkan jarak minimum dari setiap objek. Nilai yang diperoleh dalam keanggotaan data pada distance matriks adalah 0 atau 1, dimana nilai 1 untuk data yang dialokasikan ke *cluster* dan nilai 0

untuk data yang dialokasikan ke *cluster* yang lain. Lakukan perulangan dalam menentukan nilai centroid hingga nilai centroid yang dihasilkan tetap dan anggota *cluster* tidak berpindah ke *cluster* lain.

3.3.2 Algoritma Fuzzy C-means

Langkah-langkah pada algoritma Fuzzy c-means ditunjukkan pada Gambar 3.3.



Gambar 3.3 Flow chart Fuzzy c-means clustering

Langkah pertama, lakukan input data yang akan dikelompokkan (X), berupa matriks berukuran $n \times m$ (Balasko, Abonyi, & Feil, 2007). Dimana, n merupakan jumlah sampel data dan m merupakan atribut setiap data. Sehingga X_{ij} adalah data sampel ke- i ($i=1,2,\dots,n$) dan atribut ke- j ($j=1,2,\dots,m$).

Langkah kedua adalah menentukan jumlah *cluster* (c), pangkat (w), maksimum iterasi (maxIter), error terkecil yang diharapkan (ϵ), fungsi obyektif awal ($P_0 = 0$), dan iterasi awal ($t = 1$). Kemudian membangkitkan nilai acak (μ_{ik}), $i = 1, 2, \dots, n$; $k = 1, 2, \dots, c$; sebagai elemen-elemen matriks partisi awal U dan menghitung jumlah setiap kolom (atribut) menggunakan persamaan (3.7).

$$Q_j = \sum_{k=1}^c \mu_{ik} \quad (3.7)$$

Dimana Q_j adalah jumlah nilai derajat keanggotaan perkolom yang bernilai $= 1$ dengan $j=1,2,\dots,m$.

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j} \quad (3.8)$$

Langkah selanjutnya menghitung pusat *cluster* ke- k : V_{kj} , dengan $k=1,2,\dots,c$; dan $j=1,2,\dots,m$ dengan persamaan (3.9).

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (3.9)$$

dimana:

V_{kj} : pusat *cluster* ke- k untuk atribut ke- j

μ_{ik} : derajat keanggotaan untuk data sampel ke- i pada *cluster* ke- k

X_{ij} : data ke- i , atribut ke- j

Kemudian menghitung fungsi obyektif pada iterasi ke- t menggunakan persamaan (3.10) dan menghitung perubahan matriks partisi dengan persamaan (3.11).

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right) (\mu_{ik})^w \quad (3.10)$$

dimana:

V_{kj} = pusat *cluster* ke-k untuk atribut ke-j

μ_{ik} = derajat keanggotaan untuk data sampel ke-i pada *cluster* ke-k

X_{ij} = data ke-i, atribut ke-j

P_t = fungsi objektif pada iterasi ke-t

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{\frac{-1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{\frac{-1}{w-1}}} \quad (3.11)$$

dimana:

V_{kj} = pusat *cluster* ke-k untuk atribut ke-j

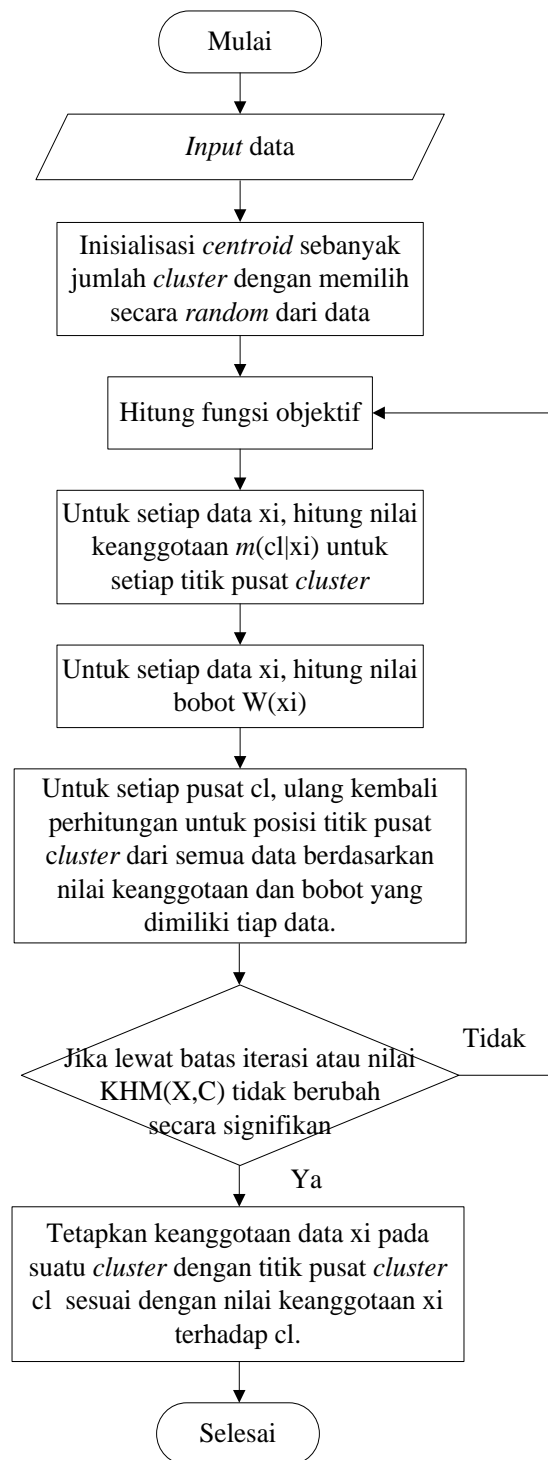
μ_{ik} = derajat keanggotaan untuk data sampel ke-i pada *cluster* ke-k

X_{ij} = data ke-i, atribut ke-j

Selanjutnya cek kondisi berhenti, jika: $(|P_t - P_{t-1}| \leq \varepsilon)$ atau $(t > \text{maxIter})$ maka proses selesai, tetapi jika tidak: $t = t + 1$, kembali lagi pada proses menghitung pusat *cluster* menggunakan persamaan (3.9).

3.3.3 Algoritma K-Harmonic Means

Adapun langkah-langkah dalam algoritma K-harmonic means ditunjukkan pada gambar 3.4.



Gambar 3.4 Flow chart K-harmonic means clustering

Pada metode ini, pertama dilakukan inisialisasi posisi titik pusat *cluster* awal secara random (Yang et al., 2009). Selanjutnya menghitung nilai fungsi objektif menggunakan persamaan (3.12), dimana p adalah input parameter. Nilai p biasanya ≥ 2 .

$$\text{KHM}(X, C) = \sum_{i=1}^N \frac{K}{\sum_{l=1}^K \frac{1}{\|x_i - c_l\|^p}} \quad 3.12$$

Untuk setiap data x_i , hitung nilai keanggotaan $m(c_l|x_i)$ untuk setiap titik pusat *cluster* c_l berdasarkan persamaan (3.13).

$$m(c_l|x_i) = \frac{\|x_i - c_l\|^{-p-2}}{\sum_{l=1}^k \|x_i - c_l\|^{-p-2}} \quad 3.13$$

Untuk setiap data x_i , hitung nilai bobot $w(x_i)$ berdasarkan persamaan (3.14).

$$w(x_i) = \frac{\sum_{l=1}^k \|x_i - c_l\|^{-p-2}}{(\sum_{l=1}^k \|x_i - c_l\|^{-p})^2} \quad 3.14$$

Untuk setiap titik pusat c_l , ulang kembali perhitungan untuk posisi titik pusat *cluster* dari semua data berdasarkan nilai keanggotaan dan bobot yang dimiliki tiap data.

$$c_l = \frac{\sum_{i=1}^N m(c_l|x_i) \cdot w(x_i) \cdot x_i}{\sum_{i=1}^N m(c_l|x_i) \cdot w(x_i)} \quad 3.15$$

Langkah selanjutnya, kembali pada proses menghitung nilai fungsi objektif hingga mendapatkan nilai yang tidak terdapat perubahan yang signifikan. Kemudian tetapkan keanggotaan data x_i pada suatu *cluster* dengan titik pusat *cluster* c_l sesuai dengan nilai keanggotaan x_i terhadap c_l . Dimana x_i merupakan anggota dari *cluster* dengan titik pusat *cluster* c_l apabila nilai keanggotaan $m(c_l|x_i)$ adalah yang terbesar dibandingkan dengan nilai keanggotaannya ke titik pusat *cluster* lain.