BAB III

FUZZY C-MEANS

3.1 *Fuzzy* Klastering

Fuzzy klastering merupakan salah satu metode analisis klaster dengan mempertimbangkan tingkat keanggotaan yang mencakup himpunan *fuzzy* sebagai dasar pembobotan bagi pengelompokan (Bezdek,1981). Metode ini merupakan pengembangan dari metode *partitioning* data dengan pembobotan *fuzzy*. Keunggulan utama *fuzzy* klastering adalah dapat memberikan hasil pengelompokan bagi objek-objek yang tersebar tidak teratur, karena jika terdapat suatu data yang penyebarannya tidak teratur maka terdapat kemungkinan suatu titik data mempunyai sifat atau karakteristik dari klaster lain. Sehingga perlu adanya pembobotan kecenderungan titik data terhadap suatu klaster. Secara matematis, masalah *fuzzy* klastering telah dirumuskan oleh Bezdek (1981) dalam bentuk optimasi kendala.

Terdapat beberapa hal yang perlu diketahui sebelum melakukan *fuzzy* klastering (Kusumadewi, 2004) :

1. Ukuran *fuzzy*

Ukuran *fuzzy* menunjukkan derajat kekaburan dari himpunan *fuzzy*. Secara umum ukuran kekaburan dapat ditulis sebagai suatu fungsi:

$$f : P(X) \rightarrow R$$

dengan P(X) adalah himpunan semua subset dari X dan *f(A)* adalah suatu fungsi yang memetakan subset A ke karakteristik derajat kekaburannya. Dalam mengukur nilai kekaburan, fungsi *f* harus mengikuti hal-hal sebagai berikut:

a. *f(A)* = 0 jika hanya jika A adalah himpunan *crisp*.

b. Jika A<B, maka *f(A)<f(B)*, di mana A<B berarti B lebih kabur dibanding A, dengan kata lain A lebih tajam dibanding B. Relasi ketajaman A<B didefinisikan dengan:

$$\mu_A[x] \le \mu_B[x], \quad \text{jika } \mu_B[x] \le 0{,}5\text{; dan} \tag{3.1}$$

$$\mu_A[x] \ge \mu_B[x], \quad \text{jika } \mu_B[x] \ge 0{,}5 \tag{3.2}$$

c. *f(A)* akan mencapai maksimum jika dan hanya jika A benar-benar kabur secara maksimum. Tergantung pada interpretasi derajat kekaburan, nilai *fuzzy* maksimal biasanya terjadi pada saat $\mu_A[x] = 0{,}5$ untuk setiap *x*.

2. Indeks Kekaburan

Indeks kekaburan adalah jarak antara suatu himpunan *fuzzy* A dengan himpunan *crisp* C yang terdekat. Himpunan *crisp* C terdekat dari himpunan *fuzzy* A dinotasikan sebagai:

$$\mu_C[x] = 0\text{, jika } \mu_A[x] \le 0{,}5\text{; dan} \tag{3.3}$$

$$\mu_C[x] = 1\text{, jika } \mu_A[x] \ge 0{,}5 \tag{3.4}$$

Ada tiga ukuran yang paling sering digunakan dalam mencari indeks kekaburan, yaitu:

a. Jarak *Hamming*

$$f(A) = \sum |\mu_A[x] - \mu_C[x]| \text{ atau} \tag{3.5}$$

$$f(A) = \sum \min[\mu_A[x], 1 - \mu_A[x]] \tag{3.6}$$

b. Jarak *Euclidean*

$$f(A) = \left\{ \sum [\mu_A[x] - \mu_C[x]]^2 \right\}^{\frac{1}{2}} \tag{3.7}$$

c. Jarak *Minkowski*

$$f(A) = \left\{ \sum \left[\mu_A[x] - \mu_C[x] \right]^m \right\}^{\frac{1}{m}} \tag{3.8}$$

3.2 *Fuzzy C-Means*

Dalam teknik klastering data, terdapat beberapa algoritma, salah satunya adalah *Fuzzy C-Means* (FCM). *Fuzzy C-Means* merupakan suatu teknik pengelompokan data di mana keberadaan tiap-tiap data dalam suatu klaster diboboti oleh derajat keanggotaan dari suatu himpunan *fuzzy* sehingga dapat mengatasi masalah tumpang tindih (*overlapping*) yang terjadi pada suatu data. Algoritma *Fuzzy C-Means* klastering pertama kali diperkenalkan oleh Dunn (1974), kemudian dikembangkan oleh Bezdek (1981), kemudian direvisi oleh Rouben (1982), Trauwert (1985), Goth dan Geva (1989), Gu dan Gubuisson (1990), Xie dan Beni (1991). Namun, algoritma FCM dari Bezdek yang paling banyak digunakan, sehingga dalam Tugas Akhir ini penulis menggunakan algoritma *Fuzzy C-Means* dari Bezdek.

Konsep dasar *Fuzzy C-Means*, pertama kali adalah menentukan pusat klaster yang akan menandai lokasi rata-rata untuk setiap klaster. Pada kondisi awal, pusat klaster ini masih belum akurat. Setiap titik data memiliki derajat keanggotaan untuk setiap klaster. Dengan cara memperbaiki pusat klaster dan derajat keanggotaan setiap titik data secara berulang, maka akan dilihat bahwa pusat klaster akan bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada minimisasi fungsi objektif yang menggambarkan jarak dari titik data yang diberikan ke pusat klaster yang terboboti oleh derajat keanggotaan titik data dari himpunan *fuzzy* tersebut.

3.2.1 Asumsi *Fuzzy C-Means*

Misalkan X suatu himpunan data berbentuk matriks berukuran $n \times p$ (n = jumlah sampel yang ada, p = banyaknya variabel) dan x_{kj} = data sampel ke-k (k = 1, 2, 3, …, n), variabel ke-j (j = 1, 2, 3, …, p) dinyatakan dalam bentuk notasi matriks sebagai berikut:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{k1} & x_{k2} & \vdots & x_{kj} & \vdots & x_{kp} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nj} & \cdots & x_{np} \end{bmatrix}.$$

Data dalam matriks X tersebut akan di kelompokkan ke dalam c (i = 1, 2, 3, …, c) buah klaster yang mempunyai derajat keanggotaan sebagai berikut.

$$\mu = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1c} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n1} & \mu_{n2} & \cdots & \mu_{nc} \end{bmatrix}$$

dan memiliki asumsi *fuzzy* klastering sebagai berikut:

1. Memiliki fungsi objektif:

$$P_t = \sum_{i=1}^{c} \sum_{k=1}^{n} (\mu_{ik})^m \left\| x_k - v_i \right\|^2 \quad (3.9)$$

di mana: x_k = observasi ke-k

v_i = pusat klaster ke-i

μ_{ik} = derajat keanggotaan himpunan *fuzzy*

c = jumlah klaster

n = jumlah observasi.

2. Memiliki nilai derajat keanggotaan untuk data ke-k di klaster ke-i adalah:

$$\mu_{ik} \in [0,1] \ , \ (1 \le i \le c; 1 \le k \le n) \quad (3.10)$$

3. Memiliki fungsi batasan (*constraint*):

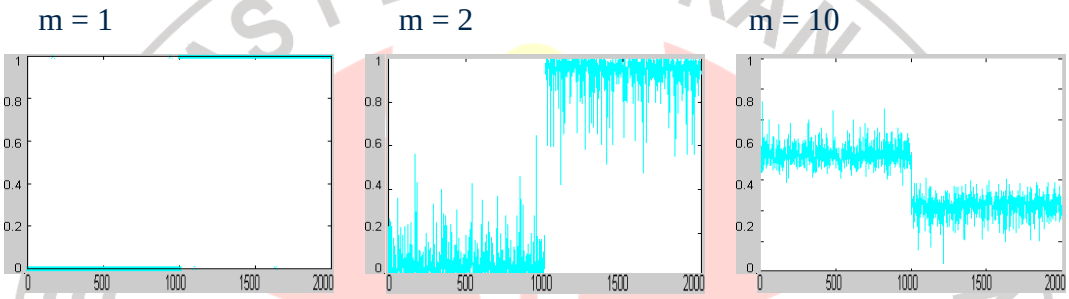
$$\sum_{i=1}^{c} \mu_{ik} = 1 \quad (3.11)$$

Sehingga $0 < \sum_{k=1}^{n} \mu_{ik} < 1$.

3.2.2 *Fuzziness Parameter*

Fuzziness parameter atau biasa disebut pangkat pembobot merupakan salah satu parameter yang diperhatikan karena berpengaruh secara signifikan pada

ke-*fuzzy*-an dari suatu hasil pengelompokkan. Indeks kekaburan ini dinotasikan sebagai *m* yang bernilai $m \in [1, \infty)$. Ketika $m \to 1$, nilai kekaburan menjadi kaku dan cenderung tegas sehingga algoritma *Fuzzy C-Means* konvergen pada perumuman k-means. Ketika $m \to \infty$, nilai kekaburan akan menjadi semakin kabur (Klir, 1995). Berdasarkan penelitian yang dilakukan oleh Klawonn (2001), nilai *m* yang sering dipakai dan dianggap yang tajam adalah $m = 2$.



Gambar 3.1 *Fuzziness Parameter*.

3.2.3 Parameter *Fuzzy C-Means*

Pada prinsipnya algoritma *Fuzzy C-Means* meminimumkan suatu fungsi objektif. Dalam meminimumkan suatu fungsi objektif diperlukan suatu metode yang dapat meminimumkan fungsi tersebut. Metode *Lagrange multiplier* (pengali *Lagrange*) biasa digunakan untuk mengoptimalkan suatu fungsi objektif yang dibatasi oleh fungsi batasan (*constraint*) dan pengali *Lagrange* yaitu λ, kemudian diturunkan terhadap parameter-parameternya dan disamakan dengan 0. Dengan *Lagrange multiplier* akan di optimumkan fungsi objektif untuk mencari parameter derajat keanggotaan dan pusat klaster (*centroid*).

Misalkan terdapat suatu fungsi yang akan dioptimumkan yaitu $f(x,y)$ dengan fungsi batasan (*constraint*) $g(x,y) = const$. Kondisi optimum dari $f(x,y)$ diperoleh pada saat $\nabla f = \lambda \nabla g$. Dengan penurunan fungsi *Lagrange* terhadap masing-masing parameter $\nabla_{x,y,\lambda} F(x,y,\lambda) = 0$, maka dapat diperoleh kondisi $\nabla f = \lambda \nabla g$. Bentuk umum *Lagrange multiplier* adalah :

$$F(x,y,\lambda) = f(x,y) + \lambda(g(x,y) - const) \tag{3.12}$$

Dalam *fuzzy* klastering, kita mempunyai $P_t = \sum_{i=1}^{c}\sum_{k=1}^{n} (\mu_{ik})^m \left\| x_k - v_i \right\|^2$ sebagai fungsi yang akan diminimumkan untuk mencari parameter μ_{ik} dan v_k. Dengan fungsi batasan yaitu $\sum_{i=1}^{c} \mu_{ik} = 1$. Maka fungsi *Lagrange* untuk FCM adalah sebagai berikut:

$$L_{FCM} = \sum_{i=1}^{c}\sum_{k=1}^{n} (\mu_{ik})^m {d_{ik}}^2 + \sum_{k=1}^{n} \lambda_k \left[\sum_{i=1}^{c} \mu_{ik} - 1 \right] \tag{3.13}$$

Kemudian akan dicari kondisi optimum untuk v_i.

$$\frac{\partial L_{FCM}}{\partial v_i} = 0 \tag{3.14}$$

$$\frac{\partial \sum_{i=1}^{c}\sum_{k=1}^{n} (\mu_{ik})^m \left\| x_k - v_i \right\|^2 + \sum_{k=1}^{n} \lambda_k \left[\sum_{i=1}^{c} \mu_{ik} - 1 \right]}{\partial v_i} = 0 \tag{3.15}$$

$$\sum_{k=1}^{n} (\mu_{ik})^m \frac{\partial \sum_{i=1}^{c} \left\| x_k - v_i \right\|^2}{\partial v_i} = 0 \tag{3.16}$$

$$-2\sum_{k=1}^{n} (\mu_{ik})^m (x_k - v_i) = 0 \tag{3.17}$$

$$-2x_k \sum_{k=1}^{n} (\mu_{ik})^m + 2v_i \sum_{k=1}^{n} (\mu_{ik})^m = 0 \tag{3.18}$$

$$2v_i \sum_{k=1}^{n} (\mu_{ik})^m = 2x_k \sum_{k=1}^{n} (\mu_{ik})^m \tag{3.19}$$

$$v_i = \frac{\sum_{k=1}^{n} x_k (\mu_{ik})^m}{\sum_{k=1}^{n} (\mu_{ik})^m} \tag{3.20}$$

Sedangkan untuk mencari kondisi optimum untuk μ_{ik} , akan lebih mudah jika dimisalkan $\|x_k - v_i\|^2 = d_{ik}^2$ sebagai berikut :

$$\frac{\partial L_{FCM}}{\partial \mu_{ik}} = 0 \tag{3.21}$$

$$\frac{\partial \sum_{i=1}^{c} \sum_{k=1}^{n} (\mu_{ik})^m \|x_k - v_i\|^2 + \sum_{k=1}^{n} \lambda_k \left[\sum_{i=1}^{c} \mu_{ik} - 1 \right]}{\partial \mu_{ik}} = 0 \tag{3.22}$$

$$m(\mu_{ik})^{m-1} d_{ik}^2 + \lambda_k = 0 \tag{3.23}$$

$$m(\mu_{ik})^{m-1} d_{ik}^2 = -\lambda_k \tag{3.24}$$

$$(\mu_{ik})^{m-1} = \frac{-\lambda_k}{m d_{ik}^2} \tag{3.25}$$

$$\mu_{ik} = \left(\frac{-\lambda_k}{m d_{ik}^2} \right)^{\frac{1}{m-1}} \tag{3.26}$$

Persamaan tersebut masih mengandung pengali *Lagrange* (λ_k), sehingga harus dibentuk sedemikian sehingga tidak mengandung pengali *Langrange* melalui fungsi batasannya.

$$\sum_{i=1}^{c} \mu_{ik} = 1 \tag{3.27}$$

$$\sum_{i=1}^{c} \mu_{ik} = \sum_{l=1}^{c} \left(\frac{-\lambda_k}{m d_{lk}^2} \right)^{\frac{1}{m-1}} = 1 \tag{3.28}$$

$$\left(\frac{-\lambda_k}{m} \right)^{\frac{1}{m-1}} = \frac{1}{\sum_{l=1}^{c} \left(\frac{1}{d_{lk}^2} \right)^{\frac{1}{m-1}}} \tag{3.29}$$

Substitusi persamaan (3.29) ke persamaan (3.26)

$$\mu_{ik} = \frac{1}{\sum_{l=1}^{c} \left(\frac{1}{d_{lk}^2} \right)^{\frac{1}{m-1}}} \cdot \left(\frac{1}{d_{ik}^2} \right)^{\frac{1}{m-1}} \tag{3.30}$$

$$\mu_{ik} = \frac{1}{\sum_{l=1}^{c} \left(\frac{d_{ik}^2}{d_{lk}^2} \right)^{\frac{1}{m-1}}} \tag{3.31}$$

3.2.4 Algoritma *Fuzzy C-Means*

Algoritma *Fuzzy C-Means* adalah sebagai berikut:

1. Input data yang akan diklaster, yaitu berupa matriks berukuran $n \times p$ (n = jumlah observasi, p = banyaknya variabel) dan x_{kj} = data observasi ke-k (k = 1, 2, 3, …, n), variabel ke-j (j = 1, 2, 3, …, p).
2. Tentukan:
   - Jumlah klaster = c;
   - *Fuzziness Parameter* = m;

- Galat (*error*) terkecil yang diharapkan = ε;
- Fungsi objektif awal = $P_0 = 0$;
- Maximum Iterasi = MaxIter.

3. Bangkitkan bilangan random μ_{ik}, *i*= 1, 2, 3, …, *c*; *k*= 1, 2, 3, …, *n*; sebagai elemen-elemen matriks partisi awal μ.
4. Hitung pusat klaster ke-*i*: V_{ij}, dengan *i*= 1, 2, 3, …, *c*; dan *j*= 1, 2, 3, …, *p*.
5. Hitung fungsi objektif pada iterasi ke-*t*, P_t.
6. Hitung perubahan matriks partisi μ_{ik}.
7. Cek kondisi berhenti:

   Jika $\left(\left|P_t - P_{t-1}\right| < \varepsilon\right)$ atau $\left(t > \text{MaxIter}\right)$ maka berhenti;

   Jika tidak, maka ulangi langkah ke-4 dengan *t* = *t*+1.

Dari algoritma tersebut dapat disimpulkan bahwa langkah pertama yang dilakukan adalah menentukan matriks derajat keanggotaan secara acak yang kemudian dijadikan acuan terhadap perhitungan pusat klaster. Pada kondisi awal, pusat klaster ini masih belum akurat, yang ditunjukkan dengan besarnya nilai selisih fungsi objektif. Sehingga dilakukan langkah iteratif dengan cara memperbaiki pusat klaster. Dengan langkah iteratif ini, dapat dilihat bahwa pusat klaster bergerak menuju lokasi yang tepat. Langkah ini dilakukan berdasarkan minimisasi fungsi objektif.

Output dari *Fuzzy C-Means* merupakan matriks pusat klaster berukuran $c \times p$ dan matriks derajat keanggotaan untuk tiap-tiap data berbentuk $n \times c$.

Pengelompokan klaster dapat dilihat dari kedua output ini. Matriks pusat klaster menunjukkan pusat klaster untuk tiap-tiap variabel yang diamati dalam setiap klasternya. Matriks derajat keanggotaan menunjukkan kecenderungan suatu data untuk masuk ke dalam klaster tertentu. Semakin besar nilai derajat keanggotaannya, maka semakin besar peluang data tersebut masuk ke dalam klaster tertentu.