

BAB II

TINJAUAN PUSTAKA DAN LANDASAN TEORI

Pada bagian ini akan memuat literatur dan landasan teori yang mendukung serta relevan dengan laporan penelitian.

2.1 Tinjauan Pustaka

Penelitian dengan menggunakan metode k-means dan metode fuzzy c-means sudah banyak dilakukan dan ada beberapa penelitian yang membandingkan kedua metode tersebut. Sehingga penelitian ini mengambil literatur-literatur untuk dijadikan bahan tinjauan pustaka dalam pemaparan dan pendukung dalam membandingkan metode k-means dan metode fuzzy c-means untuk membantu dalam menganalisa karakteristik mahasiswa berdasarkan kunjungan ke perpustakaan.

Penelitian O. J et al., (2010), Shovon & Haque, (2012) yang berjudul “*Application of k-Means Clustering algorithm for prediction of Students’ Academic Performance*” dan “*Prediction of Student Academic Performance by an Application of K-Means Clustering Algorithm*” menunjukkan bahwa *clustering* dapat digunakan untuk memonitor kinerja mahasiswa di suatu universitas. Metode ini juga dapat digunakan untuk memonitor kinerja persemester dalam meningkatkan prestasi akademik. Penelitian yang dilakukan O. J et al., (2010) menggunakan 79 data mahasiswa untuk uji coba clustering pada Universitas Nigeria, sedangkan penelitian Shovon & Haque, (2012) menggunakan 60 data mahasiswa untuk uji coba penelitiannya.

Penelitian Arora & Badal, (2013) yang berjudul “*Evaluating Student’s Performance Using k-Means Clustering*”, menggunakan algoritma K-Means karena dinilai dapat dengan cepat dan efisien membantu memantau perkembangan kinerja mahasiswa di suatu instansi pendidikan. Jumlah data yang dianalisis adalah 118 data siswa untuk mendapatkan nilai rata-rata mahasiswa tiap semester. Metode ini dapat memainkan peran penting bagi analisis akademik untuk menentukan alasan penurunan kinerja mahasiswa selama semester tertentu sehingga dapat diambil tindakan untuk meningkatkan kinerja tersebut disemester berikutnya.

Penelitian Yadav & Ahmed, (2012) yang berjudul “*Academic Performance Evaluation Using Fuzzy C-Means*”, menggunakan algoritma Fuzzy C-Means dengan jumlah data adalah sebanyak 20 data siswa untuk menganalisa dan mengetahui pemodelan kinerja akademik dan meningkatkan kualitas siswa dan guru dalam dunia pendidikan. Hasil penelitian didapat bahwa Fuzzy C-Means menjadi model terbaik untuk pemodelan kinerja akademik yang berfungsi sebagai tolak ukur dalam memantau perkembangan pemodelan siswa dalam dunia pendidikan. Diharapkan kedepannya Fuzzy C-Means dapat dikombinasikan dengan metode lain seperti Neuro Dynamic Fuzzy Expert System untuk mengevaluasi kinerja siswa dan guru serta mengembangkan sistem pembelajaran yang dapat mendukung proses pembelajaran jarak jauh berbasis internet.

Penelitian yang dilakukan Varghese et al., (2011), Inyang & Joshua, (2013) yang berjudul “*Clustering Student Data to Characterize Performance Patterns*” dan “*Fuzzy Clustering of Students Data Repository for At-Risks Students Identification and Monitoring*”, menggunakan teknik data mining dengan jumlah data 400 siswa yang dilakukan oleh Varghese et al dan 496 data siswa yang dilakukan oleh Inyang & Joshua untuk mengekstrak pola sehingga dapat dilakukan pengambilan keputusan dari database yang besar. Model Fuzzy C-Means dapat membantu dalam mencapai proses monitoring kemajuan siswa dan mengidentifikasi fitur-fitur apa saja yang mempengaruhi kinerja siswa sehingga dapat membantu para pendidik mengetahui tingkat kemampuan dari masing-masing siswa. Menurut Yusuf et al. (2012) prediksi nilai ujian akhir dapat membantu pendidik atau mahasiswa dalam melakukan tindakan ke depan sesuai dengan kondisi pada saat itu.

Penelitian yang dilakukan oleh Agbonifo & Catherine (2013) berjudul “*Fuzzy C-Means Clustering Model for Identification of Students Learning Preferences in Online Environment*” mengadopsi metode Honey dan Mumford’s learning styles dengan menggunakan teknik *clustering* Fuzzy C-Means untuk menentukan preferensi belajar dari seorang siswa dengan jumlah data yang diuji adalah sebanyak 50 data siswa sehingga membantu instruktur dan tenaga pendidik dalam proses pengajaran. Menurut Yadav & Singh (2012) dengan menggunakan metode Fuzzy C-Means dengan uji coba siswa sebanyak 20 data siswa dapat memodelkan prestasi akademik mahasiswa di lingkungan pendidikan yang digunakan sebagai

tolak ukur/dasar dalam memantau perkembangan mahasiswa pada lingkungan pendidikan. Penelitian yang dilakukan oleh Setiyono & Isnanto (2008) membahas pendekatan pengklusteran fuzzy dengan membagi kelas berdasarkan nilai prestasi mahasiswa pada mata kuliah yang menjadi persyaratan untuk menempuh mata kuliah baru dimana jumlah data yang akan dikluster adalah sebanyak 70 data mahasiswa. Hasil dari penelitian menunjukkan bahwa Fuzzy C-Means lebih cocok untuk alokasi mahasiswa menjadi beberapa kelas berdasarkan masukan beberapa jumlah kluster yang diinginkan. Munandar et al, (2013) menganalisa data mahasiswa sebanyak 126 data menurut bobot nilai mata kuliah tertentu dengan menggunakan konsep Fuzzy C-Means sehingga mampu memberikan sebuah pendukung keputusan tentang pengelompokan data nilai mahasiswa untuk menentukan klasifikasi terhadap konsentrasi yang mana seorang mahasiswa seharusnya dimasukkan.

Penelitian yang dilakukan Ghosh & Dubey (2013) berjudul “*Comparative Analysis of K-Means and Fuzzy C-Means Algorithms*”, membandingkan dua algoritma *clustering* yaitu algoritma K-Means dan algoritma Fuzzy C-Means dengan jumlah data yang diuji adalah 150 data mahasiswa. Perbandingan ini diterapkan untuk mengetahui dan mengevaluasi kinerja pengelompokan data yang paling baik. Hasil dari penelitian tersebut adalah Fuzzy C-Means menghasilkan data yang mirip dengan K-Means tetapi pengelompokan Fuzzy C-Means memerlukan waktu komputasi yang banyak dibandingkan algoritma K-Means, sehingga algoritma K-

Means menjadi algoritma yang paling baik dibandingkan algoritma Fuzzy C-Means.

2.2 Landasan Teori

2.2.1 Data Mining

Data mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa depan. Data mining adalah pengolahan data dengan skala besar, sehingga data mining memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi (Ong, 2013). Data mining juga dapat dilakukan pada berbagai jenis database dan penyimpanan informasi, namun jenis pola yang akan ditemukan ditentukan oleh berbagai fungsi data mining seperti *deskripsi class/konsep*, *asosiasi*, analisa korelasi, klasifikasi, prediksi, analisa *cluster* dan lain-lain (Dash et al., 2010).

Data mining menggunakan penerapan algoritma tertentu untuk mengekstrak pola dari data. Dimana proses ini secara otomatis akan mencari pola yang sederhana terhadap data yang besar menggunakan analisa tertentu. Data mining juga menggunakan algoritma matematika yang canggih untuk mensegmen data dan mengevaluasi kemungkinan beberapa hasil yang ditetapkan oleh pengguna (Ndehedehe et al., 2013).

Jumlah data yang dilakukan oleh suatu proses uji coba dengan menggunakan data mining adalah dapat dilakukan dengan 50 data.

2.2.2 *Clustering*

Clustering adalah proses pengelompokan benda serupa ke dalam kelompok yang berbeda, atau lebih tepatnya partisi dari sebuah data set kedalam subset, sehingga data dalam setiap subset memiliki arti yang bermanfaat. Dimana sebuah *cluster* terdiri dari kumpulan benda-benda yang mirip antara satu dengan yang lainnya dan berbeda dengan benda yang terdapat pada *cluster* lainnya. Algoritma *clustering* terdiri dari dua bagian yaitu secara *hirarkis* dan secara *partitional*. Algoritma *hirarkis* menemukan *cluster* secara berurutan dimana *cluster* ditetapkan sebelumnya, sedangkan algoritma *partitional* menentukan semua kelompok pada waktu tertentu (Madhulatha, 2012). *Clustering* juga bisa dikatakan suatu proses dimana mengelompokkan dan membagi pola data menjadi beberapa jumlah data set sehingga akan membentuk pola yang serupa dan dikelompokkan pada *cluster* yang sama dan memisahkan diri dengan membentuk pola yang berbeda ke *cluster* yang berbeda (HUNG et al., 2005).

Clustering dapat memainkan peran penting dalam kehidupan sehari-hari, karena tidak bisa lepas dengan sejumlah data yang menghasilkan informasi untuk memenuhi kebutuhan hidup. Salah satu sarana yang paling

penting dalam hubungan dengan data adalah untuk mengklasifikasikan atau mengelompokkan data tersebut ke dalam seperangkat kategori atau *cluster*. *Clustering* dapat ditemukan di beberapa aplikasi yang ada di berbagai bidang. Sebagai contoh pengelompokan data yang digunakan untuk menganalisa data statistik seperti pengelompokan untuk pembelajaran mesin, data mining, pengenalan pola, analisis citra dan *bioinformatika* (Bataineh et al., 2011).

Teknik pengelompokan saat ini dapat diklasifikasikan menjadi tiga kategori yaitu *partitional*, hirarkis dan berbasis lokalitas algoritma. Terdapat satu set objek dan kriteria *clustering* atau pengelompokan, pengelompokan *partitional* memperoleh partisi objek ke dalam *cluster* sehingga objek dalam *cluster* akan lebih mirip dengan benda-benda yang ada di dalam *cluster* dari pada objek yang terdapat pada *cluster* yang berbeda. *Partitional* mencoba untuk menguraikan dataset ke satu set *cluster* dengan menentukan jumlah *cluster* awal yang diinginkan (Varghese et al., 2011).

2.2.3 K-Means

Algoritma K-means merupakan salah satu algoritma dengan *partitional*, karena K-Means didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai *centroid* awalnya (Madhulatha, 2012). Algoritma K -means menggunakan proses secara berulang-ulang

untuk mendapatkan basis data *cluster*. Dibutuhkan jumlah *cluster* awal yang diinginkan sebagai masukan dan menghasilkan titik *centroid* akhir sebagai output. Metode K-means akan memilih pola k sebagai titik awal *centroid* secara acak atau random. Jumlah iterasi untuk mencapai *cluster centroid* akan dipengaruhi oleh calon *cluster centroid* awal secara random. Sehingga didapat cara dalam pengembangan algoritma dengan menentukan *centroid cluster* yang dilihat dari kepadatan data awal yang tinggi agar mendapatkan kinerja yang lebih tinggi (HUNG et al., 2005, Saranya & Punithavalli, 2011, Eltibi & Ashour, 2011).

Dalam penyelesaiannya, algoritma K-Means akan menghasilkan titik *centroid* yang dijadikan tujuan dari algoritma K-Means. Setelah iterasi K-Means berhenti, setiap objek dalam dataset menjadi anggota dari suatu *cluster*. Nilai *cluster* ditentukan dengan mencari seluruh objek untuk menemukan *cluster* dengan jarak terdekat ke objek. Algoritma K-means akan mengelompokkan item data dalam suatu dataset ke suatu *cluster* berdasarkan jarak terdekat (Bangoria et al., 2013). Nilai *centroid* awal yang dipilih secara acak yang menjadi titik pusat awal, akan dihitung jarak dengan semua data menggunakan rumus Euclidean Distance. Data yang memiliki jarak pendek terhadap *centroid* akan membuat sebuah *cluster*. Proses ini berkelanjutan sampai tidak terjadi perubahan pada setiap

kelompok (Agrawal & Gupta, 2013, Chaturved & Rajavat, 2013, Bhatia & Khurana, 2013).

Keuntungan Algoritma K-Means

Algoritma K-Means juga memiliki keuntungan yaitu :

1. Dalam implementasi menyelesaikan masalah, algoritma K-Means sangat *simple* serta *fleksibel*. Artinya perhitungan komputasinya tidak terlalu rumit dan algoritma ini dapat diimplementasikan pada segala bidang.
2. Algoritma K-Means sangat mudah untuk dipahami, terutama dalam implementasi data yang sangat besar serta dapat mengurangi kompleksitas data yang dimiliki (Bangoria et al., 2013)

Kelemahan Algoritma K-Means

Kelemahan yang dimiliki oleh algoritma K-Means yaitu :

1. Di Algoritma K-Means user memerlukan angka yang tepat dalam menentukan jumlah *cluster* sebanyak k karena terkadang pusat *cluster* awal dapat berubah sehingga kejadian ini bisa mengakibatkan pengelompokan data menjadi tidak stabil (Joshi & Nalwade, 2013).
2. Algoritma K-Means tidak bisa maksimal dalam menentukan atau menginisialkan nilai *centroid* awalnya, karena pada pengelompokan data dengan algoritma K-Means sangat bergantung pada nilai *centroidnya* (Ahmed & Ashour, 2011).

3. Output dari K-Means tergantung pada nilai – nilai pusat yang dipilih pada *clustering*. Sehingga pada algoritma ini nilai awal titik pusat *cluster* menjadi dasar dalam penentuan *cluster*. Pemilihan *centroid cluster* awal secara acak akan memberikan pengaruh terhadap kinerja *cluster* tersebut (Singh & Kaur, 2013; Sujatha & Sona, 2013)

Cara mengatasi kelemahan pada Algoritma K-Means

Beberapa penelitian dilakukan untuk mengatasi kelemahan yang ada pada Algoritma K-Means yaitu:

1. Kaur et al., (2013) mengusulkan perbaikan pada algoritma K-Means klasik untuk menghasilkan *cluster* yang lebih akurat. Algoritma yang diusulkan terdiri dari metode berdasarkan pemisahan data, untuk menemukan *centroid* awal sesuai dengan distribusi data. Hasil penelitian ini menunjukkan bahwa algoritma yang diusulkan menghasilkan *cluster* yang lebih baik dalam waktu perhitungan yang singkat.
2. Kodinariya & Makwana, (2013) mengusulkan beberapa cara untuk menentukan nilai k sebagai jumlah *cluster* yang dibentuk secara dinamis, salah satunya adalah dengan cara metode *elbow*. Penelitian ini menyatakan bahwa metode *elbow* akan menentukan jumlah *cluster* yang sebenarnya pada satu data set. Nilai k akan terus meningkat pada setiap langkahnya dan suatu saat nilai k akan mengalami penurunan

dengan nilai yang besar, saat seperti itulah akan terbentuk siku dari semua nilai k yang didapat dan siku tersebut menjadi nilai k yang diinginkan

Algoritma K-Means

Berikut ini langkah-langkah yang terdapat pada algoritma K-Means (Ediyanto et al., 2013)

1. Tentukan k sebagai jumlah *cluster* yang dibentuk

Untuk menentukan banyaknya *cluster* k dilakukan dengan beberapa pertimbangan seperti pertimbangan teoritis dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak *cluster*. Penelitian ini akan menggunakan metode *elbow criterion* dimana metode ini sangat praktis untuk memilih jumlah *cluster* k yang akan digunakan untuk pengelompokan data pada algoritma K-Means. (Madhulatha, 2012). Metode *elbow* ini, dapat dihasilkan dari perbandingan hasil SSE (*Sum of Squared Error*) dengan rumus SSE seperti dibawah ini (Irwanto, et. al, 2012):

$$SSE = \sum_{k=1}^K \sum_{x_i \in S_k} \|x_i - c_k\|^2 \quad (2.1)$$

Dimana X_i menyatakan norma euclid (L_2) dan C_k adalah pusat kluster S_k yang dihitung berdasarkan rata-rata jarak titik-titik kluster ke pusat kluster.

2. Tentukan k *centroid* (titik pusat *cluster*) awal secara random

Penentuan *centroid* awal dilakukan secara random/acak dari objek-objek yang tersedia sebanyak k *cluster*, kemudian untuk menghitung *centroid cluster* ke- i berikutnya, digunakan rumus sebagai berikut :

$$v = \frac{\sum_{i=1}^n x_i}{n} \quad ; i = 1, 2, 3, \dots, n \quad (2.2)$$

Dimana; v : *centroid* pada *cluster*

x_i : objek ke- i

n : banyaknya objek/jumlah objek yang menjadi anggota *cluster*

3. Hitung jarak setiap objek ke masing-masing *centroid* dari masing-masing *cluster*. Untuk menghitung jarak antara objek dengan *centroid* dapat menggunakan Euclidian Distance

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad ; i = 1, 2, 3, \dots, n \quad (2.3)$$

Dimana; x_i : objek x ke- i

y_i : daya y ke- i

n : banyaknya objek

4. Alokasikan masing-masing objek ke dalam *centroid* yang paling dekat. Untuk melakukan pengalokasian objek kedalam masing-masing *cluster* pada saat iterasi secara umum dapat dilakukan dengan cara hard k-means dimana secara tegas setiap objek dinyatakan sebagai anggota

cluster dengan mengukur jarak kedekatan sifatnya terhadap titik pusat *cluster* tersebut.

5. Lakukan iterasi, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan (2.2)
6. Ulangi langkah 3 jika posisi *centroid* baru tidak sama

2.2.4 Fuzzy C-Means

Algoritma Fuzzy C-Means diusulkan pertama kali oleh Dunn pada tahun 1973 dan kemudian diperbaharui oleh Bezdek pada tahun 1981. Algoritma ini merupakan salah satu teknik *soft clustering* yang paling populer dengan menggunakan pendekatan data *point* dimana titik pusat *cluster* akan selalu diperbaharui sesuai dengan nilai keanggotaan dari data yang ada dan selain itu algoritma fuzzy c-means juga merupakan algoritma yang bekerja dengan menggunakan model fuzzy sehingga memungkinkan semua data dari semua anggota kelompok terbentuk dengan derajat keanggotaan yang berbeda antara 0 dan 1 (Bora & Gupta, 2014; Sanmorino, 2012). Metode Fuzzy C-Means pada dasarnya memiliki tujuan meminimalisasikan fungsi serta mendapatkan pusat *cluster* yang nantinya akan digunakan untuk mengetahui data yang masuk ke dalam sebuah *cluster* (Simbolon et al., 2013).

Algoritma Fuzzy C-Means sudah banyak dipergunakan pada berbagai aplikasi seperti pada penambangan data, obat-obatan, pencitraan, deteksi

pola, bioinformatika dan aplikasi ilmiah serta rekayasa lainnya. Salah satu tujuan dari algoritma ini adalah mengelompokkan data set agar menghasilkan informasi yang efisien dan akurat (Grover, 2014). Karakteristik algoritma Fuzzy C-Means adalah dapat mengelompokkan data yang kabur dan menghasilkan pengelompokan data dimana bobot keanggotaannya berasal dari data tersebut (Ghosh & Dubey, 2013).

Fuzzy C-Means berhubungan dengan konsep kesamaan fungsi objek yang berdekatan dan menemukan titik pusat *cluster* sebagai *prototype*. Untuk beberapa objek data tidak memiliki batasan pada salah satu kelas saja tetapi data tersebut dapat dikelompokkan berdasarkan derajat keanggotaan yaitu antara 0 dan 1 yang menunjukkan keanggotaan parsial dari data tersebut (Phukon & Baruah, 2013). Beberapa contoh dalam penerapan Fuzzy C-Means adalah masalah pengelompokan data nyata yang telah dibuktikan dengan menghasilkan karakteristik data yang baik (Phukon & Baruah, 2013). Algoritma ini dimulai dengan menentukan jumlah *cluster* yang diinginkan serta menginisialisasikan nilai keanggotaan yang berisikan semua data kemudian akan dikelompokkan berdasarkan *clusternya*. Pusat-pusat *cluster* dihitung dari jarak terdekat ke titik-titik yang memiliki nilai keanggotaan lebih besar. Dengan kata lain, nilai-nilai keanggotaan tersebut akan bertindak sebagai nilai bobot sementara pada suatu *cluster* (K.G & Patnaik, 2006).

Keuntungan Algoritma Fuzzy C-Means

Algoritma Fuzzy C-Means memiliki keuntungan yaitu

1. Dalam implementasi menyelesaikan masalah algoritma Fuzzy C-Means dapat memahami karakteristik data yang kabur atau data yang tidak terdefinisikan.
2. Memiliki kemampuan dalam mengelompokan data yang besar
3. Lebih kokoh terhadap data *outlier*/data dengan karakter yang berbeda atau *value* yang berbeda dalam satu atau beberapa variabel
4. Penentuan titik *cluster* yang optimal (Ali et al., 2008; Suganya & Shanthi, 2012; Martino & Sessa, 2009)
5. Dapat melakukan *clustering* lebih dari satu variabel secara sekaligus (Simbolon et al., 2013).

Kelemahan Algoritma Fuzzy C-Means

Kelemahan yang dimiliki oleh algoritma Fuzzy C-Means yaitu :

1. Pada algoritma Fuzzy C-Means user memerlukan lebih banyak waktu untuk proses perhitungan komputasinya dalam menentukan *cluster* pada setiap anggota di suatu dataset (Bora & Gupta, 2014)
2. Masih terpengaruh terhadap cara pembagian data yang sering dipergunakan pada data yang sama dan sangat sensitif terhadap kondisi awal seperti jumlah *cluster* dan titik pusat *cluster* pada pengelompokan data (Lu et al., 2013).

Cara mengatasi kelemahan pada Algoritma Fuzzy C-Means

Beberapa penelitian dilakukan untuk mengatasi kelemahan yang ada pada algoritma Fuzzy C-Means yaitu :

1. (Pravitasari, 2009), mengusulkan perbaikan pada algoritma Fuzzy C-Means yaitu meminimalkan objektif pengelompokan data dengan salah satu parameternya adalah fungsi keanggotaan dalam fuzzy. Untuk mengurangi lamanya perhitungan/komputasi pada pengelompokan data yang terjadi pada algoritma Fuzzy C-Means, penelitian ini menggunakan proporsi *eigen value* dari kemiripan data yang akan dikelompokan (Lu et al., 2013).
2. Lu et al., (2013), memiliki usul mengintegrasikan algoritma Fuzzy C-Means dengan WEKA (Waikato Environment for Knowledge Analysis) dimana akan memiliki sedikit jumlah iterasi. Dan selain itu akan memberikan cara dalam menentukan titik pusat *cluster* awal agar sesuai dengan karakteristik data tersebut sehingga titik pusat *cluster* yang sudah terbentuk tidak sensitif terhadap kondisi awal.

Algoritma Fuzzy C-Means

Langkah-langkah algoritma Fuzzy C-Means (Andriyani et al., 2013) yaitu :

1. Masukkan data yang akan di*cluster* berupa matriks X berukuran $n \times m$ (n =banyaknya sampel data dan m =banyaknya variabel setiap data).
 X_{ij} =data sampel ke- i ($i=1,2,\dots,n$), variabel ke- j ($j=1,2,\dots,m$).

2. Tentukan :

- Banyaknya kluster yang akan dibentuk (c)
- Pangkat pembobot (w)
- Maksimum iterasi ($MaxItr$).
- Error terkecil (ξ)
- Fungsi obyektif awal ($P_0=0$)
- Iterasi awal ($t=1$)

3. Bangkitkan matriks partisi awal $U_{n \times c} = [\mu_{ik}]$, μ_{ik} yaitu bilangan random yang menyatakan suatu derajat keanggotaan

4. Hitung pusat *cluster* ke- k (v_{kj}) dengan $k=1,2,\dots,c$; dan $j=1,2,\dots,m$ sebagai berikut :

$$V_{kj} = \frac{\sum_{i=1}^n (\mu_{ik})^w \cdot X_{ij}}{\sum_{i=1}^n (\mu_{ik})^w} \quad (2.4)$$

5. Hitung fungsi obyektif pada iterasi ke- t , P_t , yang menggambarkan jumlah jarak data ke pusat *cluster*.

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (2.5)$$

Dengan

P_t = fungsi obyektif;

X_{ij} = elemen X baris i , kolom j ;

V_{kj} = pusat *cluster*

6. Perbaiki derajat keanggotaan matriks partisi :

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-1}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-1}} \quad (2.6)$$

dengan :

$i = 1, 2, \dots, n$

$k = 1, 2, \dots, c$

X_{ij} = sampel data ke- i , variabel ke- j

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

w = pangkat pembobotan

7. Cek kondisi berhenti

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

- Jika tidak: $t=t+1$, ulangi langkah ke-4