

BAB II

KAJIAN PUSTAKA DAN DASAR TEORI

2.1 Perilaku Konsumen

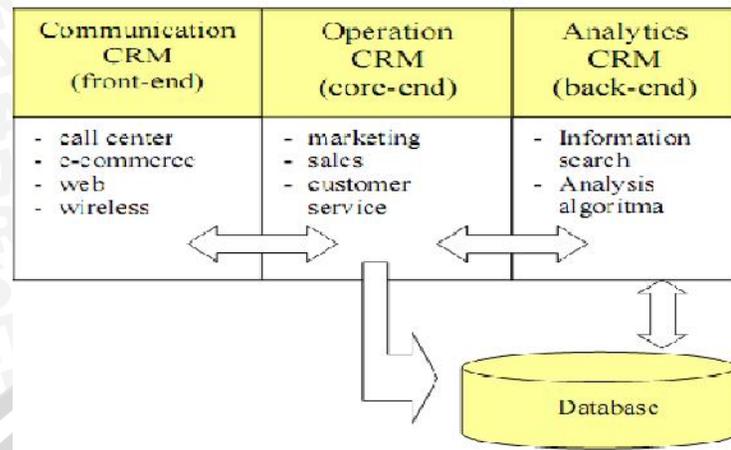
Perilaku konsumen adalah perilaku yang diperhatikan konsumen dalam mencari, membeli, menggunakan, mengevaluasi dan mengabaikan produk, jasa, atau ide yang diharapkan dapat memuaskan konsumen dalam mengkonsumsi produk atau jasa yang ditawarkan.

Pelayanan kepada konsumen bertujuan memelihara dan meningkatkan hubungan psikologis antara produsen dan pelanggan serta memantau berbagai keluhan pelanggan. Jika dilihat dari daya tarik transaksinya, pelanggan terbagi dalam empat kategori yakni *star customer*, *question mark customer*, *profit making customer* dan *profit taking customer*. Tujuan mengidentifikasi pelanggan adalah untuk mengetahui pelanggan yang bernilai dan pelanggan yang tidak bernilai. Menurutnya pelanggan yang potensial tidak boleh dibiarkan, tapi harus dikelola dan diperhatikan lewat berbagai pelayanan [KER-07].

Potensi pelanggan adalah nilai-nilai positif yang dimiliki pelanggan. Potensi pelanggan sangat erat kaitannya dengan karakter pelanggan. Karakter pelanggan bisa diketahui lewat perilaku pelanggan. Perilaku pelanggan adalah interaksi dinamis antara pengaruh dan kognisi, perilaku dan kejadian di sekitar kita dimana manusia melakukan aspek pertukaran dalam hidup mereka [RAN-06].

2.1.1 Manajemen Hubungan Pelanggan

Manajemen hubungan pelanggan (*customer relationship management* atau CRM) adalah sebuah terminologi yang menggambarkan bagaimana kita berinteraksi dan proaktif mengatur hubungan pelanggan kita. CRM merupakan sistem nyata bagaimana kita bekerja dengan para pelanggan, memecahkan problem pelanggan, mendorong pelanggan untuk membeli produk dan jasa serta melakukan kesepakatan dalam transaksi finansial. CRM katanya lagi merupakan sebuah sistem dalam proses bisnis, sebuah teknologi dan seperangkat aturan untuk melakukan kesepakatan dengan pelanggan pada berbagai tingkatan dalam aktifitas bisnis [CUN-02].



Gambar 2.1 Gambar Generalisasi Arsitektur sistem CRM
(Sumber : Zumenstain, 2007)

Ada empat tahap yang harus dilakukan CRM agar fungsinya benar-benar bisa berjalan yakni identifikasi pelanggan, diferensiasi pelanggan, interaksi dan customized. Identifikasi pelanggan lanjutnya bisa dimulai dengan melakukan pengumpulan data seperti nama, alamat, nomor telepon, jabatan lainnya. Langkah berikutnya adalah diferensiasi yakni mengkategorikan pelanggan menurut nilai dan kebutuhannya. Hal ini bertujuan untuk mengukur sejauh mana keuntungan total yang bisa diperoleh perusahaan apabila melanjutkan transaksi dengan pelanggan. Langkah selanjutnya adalah menentukan nilai-nilai yang dimiliki pelanggan. Apakah menguntungkan bagi perusahaan atau tidak. Setelah semua tersebut diketahui maka perusahaan bisa melakukan kustomisasi pelayanan yang tepat diberikan kepada pelanggan [KER-07].

2.2 Tarif Dasar Listrik

2.2.1 Pengertian

Tarif dasar listrik atau biasa disingkat TDL, adalah tarif yang boleh dikenakan oleh pemerintah untuk para pelanggan. PLN adalah satu-satunya perusahaan yang boleh menjual listrik secara langsung kepada masyarakat Indonesia, maka TDL bisa dibidang adalah tarif untuk penggunaan listrik di Indonesia. Saat ini TDL rata-rata adalah. Pada 2004, tarif nonsubsidi pelanggan 6.600 VA ke atas sekitar Rp 1.380 per kilowatt-hour (kWh), sedang tarif subsidi sekitar Rp 600 per kWh. Pada awal 2008,

diberlakukan tarif non subsidi untuk pelanggan listrik dengan daya 6600 keatas.

Mulai tahun 2010, pemerintah memutuskan menaikkan TDL rata-rata 10%. Hal ini didasarkan pada Pasal 8 UU No.2 Tahun 2010, untuk menutupi kekurangan subsidi sebesar Rp4,8 triliun karena alokasi anggaran subsidi listrik ditetapkan Rp.55,1 triliun. Tetapi untuk TDL 450-900 VA, memutuskan tidak ada kenaikan [TDL-10].

2.2.2 Pembagian Tarif Berdasarkan Kelompok Pengguna

Berdasarkan persetujuan DPR RI tersebut dan mengacu pada Pasal 34 Undang-Undang Nomor 30 Tahun 2009 tentang Ketenagalistrikan, Pemerintah telah menerbitkan Peraturan Menteri Energi dan Sumber Daya Mineral Nomor 07 Tahun 2010 tanggal 30 Juni 2010 tentang Tarif Tenaga Listrik yang disediakan oleh Perusahaan Perseroan (Persero) PT Perusahaan Listrik Negara (Berita Negara RI Tahun 2010 No. 314) yang memuat Struktur maupun penggolongan tarifnya tidak mengalami perubahan. Penggolongan tarif tersebut didasarkan pada penggunaan listrik itu sendiri. Adapun penggolongannya dibagi menjadi golongan tarif sosial, golongan tarif rumah tangga, golongan tarif bisnis dan industri. Dalam skripsi ini, data golongan yang digunakan adalah golongan tarif bisnis dan golongan industri [TDL-10].

2.2.2.1 Golongan Tarif Bisnis

Pelanggan yang termasuk kedalam golongan tarif Bisnis adalah Pelanggan yang sebagian atau seluruh tenaga listrik dari PT PLN (Persero) digunakan untuk salah satu atau beberapa kegiatan berbentuk [TDL-10]:

1. Usaha jual beli barang, jasa, dan perhotelan.
2. Usaha perbankan.
3. Usaha perdagangan ekspor/impor.
4. Kantor Firma, CV, PT atau badan hukum/perorangan yg bergerak dalam bidang usaha perdagangan.
5. Usaha pergudangan dimana sebagian atau seluruh bangunan digunakan untuk tempat penyimpanan badang atau material.

6. Usaha peorangan atau badan hukum yang sebagian besar atau seluruh kegiatannya merupakan penjualan barang atau jasa.
7. Usaha-usaha lainnya yang bertendensi komersial seperti praktek dokter, dan lain sebagainya.

Sebagai contoh adalah hotel dan apartemen. Tarif Dasar Listrik untuk keperluan Bisnis, terdiri atas [TDL-10]:

1. Golongan tarif untuk keperluan bisnis kecil pada tegangan rendah, dengan daya 450 VA s.d, 5.500 VA (B-1/TR).
2. Golongan tarif untuk keperluan bisnis menengah pada tegangan rendah, dengan daya 6.600 VA s.d, 200 kVA (B-2/TR).
3. Golongan tarif untuk keperluan bisnis besar pada tegangan menengah, dengan daya di atas 200 kVA (B-3/TM).

2.2.2.2 Golongan Tarif Industri

Pelanggan yang termasuk kedalam golongan tarif Industri adalah pelanggan yang kegiatan usahanya berupa industri pabrik produksi barang, sebagai contoh adalah pabrik. Tarif Dasar Listrik untuk keperluan Industri, terdiri atas [TDL-10]:

1. Golongan tarif untuk keperluan industri kecil/industri rumah tangga pada tegangan rendah, dengan daya 450 VA s.d. 14 kVA (I-1/TR).
2. Golongan tarif untuk keperluan industri sedang pada tegangan rendah, dengan daya di atas 14 kVA s.d, 200 kVA (I-2/TR).
3. Golongan tarif untuk keperluan industri menengah pada tegangan menengah, dengan daya di atas 200 kVA (I-3/TM).
4. Golongan tarif untuk keperluan industri besar pada tegangan tinggi, dengan daya 30.000 kVA ke atas (I-4/TT).

2.3 Data Clustering

2.3.1 Definisi Clustering

Clustering adalah metode analisis data, yang sering dimasukkan sebagai salah satu metode *Data mining*, yang tujuannya adalah untuk mengelompokkan data dengan karakteristik yang sama ke suatu ‘wilayah’ yang sama dan data dengan karakteristik

yang berbeda ke ‘wilayah’ yang lain (SHI-00). Langkah awal dari metode ini adalah dengan mendapatkan subset–subset dari objek–objek yang terhubung untuk kemudian mencari definisi dari masing–masing subset. *Clustering* dan segmentasi sebenarnya mempartisi *database*, karena itu setiap partisi atau *group* adalah sama menurut kriteria atau metrik tertentu. Jika pengukuran kesamaan tersedia, maka terdapat sejumlah teknik untuk membentuk *cluster*.

Analisis *cluster* dapat diterapkan pada bidang apa saja. Namun pemakaian teknik ini lebih familiar pada bidang pemasaran karena memang salah satu kegiatan yang dilakukan dalam pemasaran adalah pengelompokan, yang disebut segmentasi pasar. Tujuan analisis *cluster* di dalam pemasaran adalah sebagai berikut [HK-01]:

- 1. Membuat segmen pasar (*segmenting the market*)**

Pelanggan atau pembeli sering diklasterkan berdasarkan manfaat atau keuntungan yang diperoleh dari pembelian barang. Setiap *cluster* akan terdiri dari pelanggan/pembeli yang relatif homogen, dinyatakan dalam manfaat yang dicari.

- 2. Memahami perilaku pembeli**

Analisis *cluster* digunakan untuk mengenali atau mengidentifikasi kelompok pembeli yang homogen. Kemudian perilaku dalam untuk setiap kelompok perlu dikaji secara terpisah. Responden (pembeli) dikelompokkan didasarkan pada *self-reported importance* yang terkait pada setiap faktor pilihan yang digunakan untuk memilih pasar atau *mall* di mana para pembeli membeli barang yang dibutuhkan.

- 3. Mengenali peluang produk baru**

Dengan mengklasterkan merek dan produk, *competitive set* di dalam pasar bisa ditentukan. Merek di dalam *cluster* yang sama bersaing sengit satu sama lain, daripada merek dari *cluster* lain.

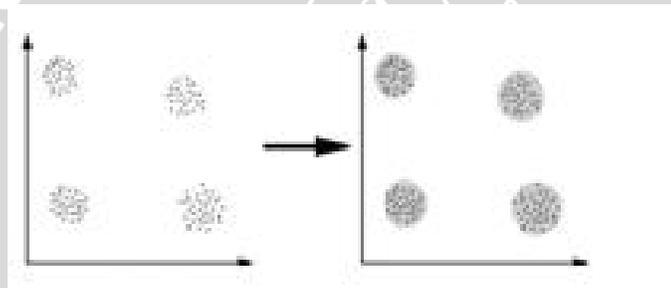
- 4. Mereduksi data**

Analisis *cluster* digunakan sebagai suatu alat mereduksi data secara umum, untuk mengembangkan *cluster* atau *subgroup* dari data yang mudah dikelola dari kumpulan data asli, secara individual.

2.3.2 Data Clustering

Data *clustering* adalah suatu metode pemecahan masalah dengan memodelkan suatu bentuk masalah dalam model matematika dengan mencari bentuk (*structure*) menggunakan persamaan sifat yang ada pada data tersebut. Proses terpenting dalam *clustering* ini adalah melihat persamaan yang terdapat pada data, sehingga dapat digolongkan pada kelas-kelas (*cluster*) yang ditentukan. Cara kerja data *clustering* adalah mengkoordinasi data-data yang ada ke dalam beberapa kelas, yang anggota dari kelas tersebut memiliki kesamaan dalam hal-hal tertentu.

Sebuah *cluster* adalah kumpulan data-data yang memiliki kesamaan diantara data-data itu sendiri dan ketidaksamaan pada data-data yang terdapat pada kelas lain. Contoh *clustering* dapat dideskripsikan dalam gambar 2.2



Gambar 2.2 Data Clustering Dengan Hard Cluster

Sumber: *a Tutorial On Clustering Algorithm*. 2007

http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/

Jenis-jenis *clustering* berdasar kesamaan pada data:

1. *Distance Based Clustering*

Dalam gambar 2.2, data terbagi atas 4 *cluster*. Kriteria kesamaan pada data tersebut adalah jarak (*distance*). Jika dua atau lebih data yang memiliki jarak yang berdekatan berdasar sebuah jarak yang ditentukan (dalam hal ini jarak geometris) maka hal ini disebut *distance-based clustering*.

2. *Conceptual Clustering*

Jika dua atau lebih data yang di miliki oleh sebuah *cluster* memiliki kesamaan umum dalam konsep tertentu dan bukan berdasarkan kesamaan pengukuran, hal ini disebut sebagai *conceptual clustering*.

Data *clustering* yang secara umum digunakan untuk pengolahan data matematis adalah *distance-based clustering*, karena *clustering* jenis ini dilakukan secara matematis. *Fuzzy clustering* termasuk dalam *distance based clustering*.

Distance based clustering memiliki beberapa jenis algoritma yang umum digunakan, yaitu :

a. *Exclusive Clustering*

Clustering jenis ini biasa disebut dengan *hard clustering*, karena algoritma *clustering* jenis ini menentukan apakah data tersebut masuk pada *cluster* tertentu dengan pernyataan ya atau tidak (0,1). Salah satu algoritma *clustering* jenis ini adalah *K-Means*.

b. *Overlapping Clustering*

Algoritma *clustering* jenis ini menentukan seberapa besar suatu data menjadi milik *cluster* tertentu (0,1). *Clustering fuzzy C-Means* termasuk dalam algoritma jenis ini.

c. *Hierarchichal Clustering*

Algoritma *clustering* jenis ini menggunakan sistem *cluster* yang bertingkat dengan cara menggabungkan *cluster-cluster* terdekat.

d. *Probabilistic Clustering*

Algoritma *clustering* ini menggunakan sebuah model seperti *Gaussian* atau yang lainnya sebagai *cluster*.

Dalam algoritma *distance based clustering* yang memegang peranan terbesar adalah *distance measure* atau pengukuran jarak. *Distance* atau jarak memiliki kemampuan yang akan menentukan sebuah data atau sebuah obyek akan masuk ke dalam kelas yang mana.

2.4 System Fuzzy

Terdapat beberapa hal yang perlu diketahui dalam memahami sistem *fuzzy*, yaitu [KUS-03]:

1. Peubah *Fuzzy*

Merupakan peubah yang akan dibahas dalam suatu sistem *fuzzy*. Contoh: umur, temperature, permintaan, dsb.

2. Himpunan *fuzzy*

Merupakan suatu kelompok yang mewakili suatu kondisi atau keadaan tertentu dalam suatu peubah *fuzzy*. Contoh: peubah umur terbagi menjadi 3 himpunan *fuzzy*, yaitu muda, parobaya dan tua. Himpunan *fuzzy* didasarkan pada gagasan untuk memperluas jangkauan fungsi karakteristik sedemikian hingga fungsi tersebut akan mencakup bilangan real pada interval [0,1]. Derajat keanggotaan menunjukkan bahwa suatu objek data dalam semesta pembicaraan tidak hanya berada pada 0 atau 1, namun juga berda pada nilai yang terletak di anantara 0 sampai 1. Dengan kata lain, derajat keanggotaan suatu objek data tidak hanya bernilai benar atau salah. Nilai 0 menunjukkan salah, nilai 1 menunjukkan benar, dan masih ada nilai-nilai yang terletak di antara benar dan salah.

Terdapat beberapa operasi himpunan *fuzzy* yang didefinisikan secara khusus untuk mengkombinasikan dan memodifikasi himpunan *fuzzy*, diantaranya [KUS-02]:

a. Interseksi

Operator ini ditunjukkan dengan derajat keanggotaan () minimum objek data antar kedua kelompok yang didefinisikan pada persamaan 2.1.

$$\mu_{A \cap B} = \min(\mu_A[x], \mu_B[y]) \quad (2.1)$$

b. Union

Operator ini ditunjukkan dengan derajat keanggotaan () maksimum objek data antar kedua kelompok yang didefinisikan pada persamaan 2.2.

$$\mu_{A \cup B} = \max(\mu_A[x], \mu_B[y]) \quad (2.2)$$

c. Komplement

Komplement atau negasi suatu peubah A berisi semua objek data yang tidak berada pada peubah A, didefinisikan pada persamaan 2.3.

$$\mu_{A^c}[x] = 1 - \mu_A[x] \quad (2.3)$$

3. Fungsi keanggotaan

Fungsi keanggotaan merupakan suatu kurva yang menunjukkan pemetaan objek data ke dalam derajat keanggotaan yang memiliki interval antara 0 sampai 1. Sedangkan tinggi himpunan *fuzzy* adalah derajat keanggotaan maksimum suatu objek data.

Dalam algoritma *fuzzy c-means* nilai derajat keanggotaan didapatkan antara 0 sampai 1. Dalam kasus ini garis fungsi keanggotaannya membentuk garis yang lebih halus untuk mengindikasikan bahwa setiap objek titik data mungkin dapat menjadi kelompok lain dengan derajat keanggotaan berbeda.

2.5 Fuzzy Clustering

Pada analisis *cluster* klasik, kelompok-kelompok yang terbentuk setelah proses *clustering*, harus merupakan pembagian data set sedemikian rupa, sehingga untuk data yang berada pada kelompok lain. Walaupun demikian, syarat ini seringkali dirasakan terlalu kaku dan sulit dipenuhi pada aplikasi praktek, sehingga perlu diganti dengan syarat yang lebih ringan. Untuk itu digunakanlah kompromi dengan prinsip-prinsip *fuzzy* yang kemudian dikenal sebagai *fuzzy clustering*.

Dengan demikian dapat dikatakan *fuzzy clustering* adalah suatu teknik untuk mengelompokkan data, yaitu membagi *dataset* yang diberikan ke dalam kelas-kelas atau kelompok-kelompok yang dikenal dengan sebutan *cluster*. Sedangkan tujuan dari pembagian ini adalah untuk memisahkan *dataset* sedemikian rupa hingga untuk dua buah data pada *cluster* yang sama adalah semirip mungkin, dan untuk dua data pada *cluster* yang tidak sama adalah seberbeda mungkin. Penggolongan data ke dalam *cluster* ditentukan berdasar derajat keanggotaan (*membership degree*) yang bernilai antara 0 dan 1, dengan demikian dimungkinkan adanya sebuah data berada pada beberapa *cluster* sekaligus dalam waktu yang sama. Derajat keanggotaan ini

umumnya ditentukan berdasarkan jarak data terhadap masing-masing pusat *cluster*. Bila dalam observasi memiliki variabel, dan k adalah jumlah *sampel training*, maka Z_s dapat didefinisikan dalam persamaan 2.4.

$$Z_s = [Z_{1s}, Z_{2s}, Z_{3s}, \dots, Z_{ns}]^T, Z_s \in \mathbb{R}^n \quad (2.4)$$

Bila terdapat N observasi, maka Z dinyatakan dalam persamaan 2.5.

$$Z = \{Z_s | s=1, 2, \dots, N\} \quad (2.5)$$

Dan direpresentasikan oleh $n \times N$ pada persamaan 2.6.

$$Z = \begin{bmatrix} Z_{11} & \dots & Z_{1N} \\ \vdots & \ddots & \vdots \\ Z_{n1} & \dots & Z_{nN} \end{bmatrix} \quad (2.6)$$

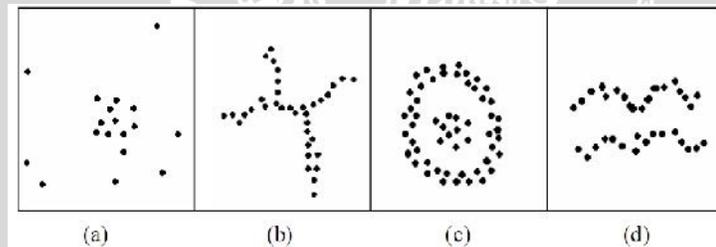
Z biasa disebut matriks data. Kolom dan baris disebut *features* (atau *atributes*) dan *pattern* (atau *object*). Dalam matriks, biasanya kesamaan didefinisikan sebagai suatu jarak dari vektor data ke sebuah objek atau pusat *cluster*. Pusat dari *cluster* biasanya belum diketahui sebelumnya dan dicari menggunakan algoritma *clustering* pada saat membuat partisi [KUS-02].

2.6 Fuzzy C-Means

Fuzzyc-means pertamakali dikemukakan oleh Dunn dan dikembangkan oleh Bezdek yang banyak digunakan dalam *pattern recognition*. Metode ini merupakan pengembangan dari metode *nonhierarki K-Means Cluster*, karena pada awalnya ditentukan dulu jumlah kelompok atau *cluster* yang akan di bentuk. Kemudian dilakukan iterasi sampai mendapatkan keanggotaan kelompok tersebut. Pemilihan metode ini didasarkan pada beberapa jurnal dan penelitian sebelumnya yang mengindikasikan bahwa metode *fuzzy cluster* merupakan metode yang paling robust, karena pusat *cluster* dan hasil pengelompokkan tidak berubah jika ada data baru yang ekstrim [KLW-00].

Metode ini juga memberikan hasil yang *smooth* (halus) karena pembobotan yang digunakan berdasarkan himpunan *fuzzy* [SHI-00]. Kehalusan disini berarti objek pengamatan tidak mutlak untuk menjadi anggota satu kelompok saja, tapi juga mungkin menjadi anggota kelompok yang lain dengan ukuran tingkat keanggotaan yang berbeda-beda. Objek akan cenderung menjadi anggota kelompok tertentu dimana tingkat keanggotaan objek dalam kelompok itu paling besar dibandingkan dengan kelompok lainnya.

FCM baik digunakan untuk mengelompokkan objek terutama jika objek-objek tersebut tersebar berserakan dan terdapat nilai ekstrim didalamnya. Ketidakteraturan bukan berarti objek-objek tersebut tidak berpola, namun yang dimaksud ketidakteraturan ini berarti tidak ada kecenderungan yang pasti bahwa objek-objek tersebut akan mengelompok secara jelas. Pada gambar 2.3, diperlihatkan beberapa jenis sebaran objek yang mengandung nilai ekstrim (a) dan sebaran berpola namun sulit ditentukan pengelompokkannya secara jelas (b), (c), dan (d).



Gambar2.3 Gambar sebaran objek

(Sumber : Shihab, 2000)

Salah satu penerapan logika *fuzzy* adalah dalam *clustering* atau pengelompokan. *Fuzzy clustering* adalah bagian dari *pattern recognition* atau pengenalan pola. *Fuzzy clustering* memainkan peran yang paling penting dalam pencarian struktur dalam data. *Fuzzy clustering* adalah salah satu teknik untuk menentukan *cluster* optimal dalam suatu ruang vektor yang didasarkan pada bentuk normal *Euclidian* untuk jarak antar vektor [KUS-04].

Ada dua metode dasar dalam *fuzzy clustering*. Metode pertama disebut dengan *fuzzy c-means*. Metode ini dinamakan demikian karena dengan *clustering* ini

akan dibentuk sebanyak c -cluster yang sudah ditentukan sebelumnya. Metode yang kedua adalah metode yang banyaknya $cluster$ tidak ditentukan sebelumnya. Metode ini dinamakan dengan *fuzzy subtractive clustering* [KUS-04]. Metode *clustering* yang akan digunakan dalam penelitian ini adalah *fuzzy c-means*. Metode ini pertama kali dikenalkan oleh Jim Bezdek pada tahun 1981 [KUS-04]. *Fuzzy c-means* adalah salah satu teknik pengklusteran data yang mana keberadaan tiap titik data dalam suatu $cluster$ ditentukan oleh derajat keanggotaan. Algoritma dari *fuzzy c-means* adalah sebagai berikut.

FCM merupakan algoritma yang bersifat iteratif. Tujuan dari *FCM* adalah untuk menemukan pusat $cluster$ (*centroid*) yang meminimalkan fungsi pertidaksamaan. Berikut langkah-langkah dari *FCM* [KUS-04]:

1. Input data yang akan dicluster X , berupa matrix berukuran $n \times m$ (n = jumlah sampel data, m = atribut setiap data).
 X_{ij} data sampel ke- i ($i = 1, 2, \dots, n$), atribut ke- j ($j = 1, 2, \dots, m$)
2. Menentukan jumlah $cluster$ (c), pangkat (w), maksimum iterasi (MaxIter), error terkecil yang diharapkan (ξ), fungsi objektif awal ($P_0 = 0$), dan iterasi awal ($t = 1$).
3. Membangkitkan bilangan random η_{jk} , $j = 1, 2, \dots, n$, $k = 1, 2, \dots, c$ sebagai elemen matriks partisi awal U .

Menghitung jumlah setiap kolom (atribut) Dengan $i = 1, 2, \dots, n$ dengan persamaan 2.7 dan 2.8.

$$Q_j = \sum_{k=1}^c \mu_{jk} \quad (2.7)$$

$$\mu_{ik} = \frac{\mu_{jk}}{Q_j} \quad (2.8)$$

4. Menghitung pusat $cluster$ ke- k : V_{kj} , dengan $k = 1, 2, \dots, c$; dan $j = 1, 2, \dots, m$. dengan persamaan 2.9.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (2.9)$$

5. Menghitung fungsi obyektif pada iterasi ke- t , dengan persamaan 2.10.

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (2.10)$$

6. Menghitung perubahan matriks partisi dengan persamaan 2.11.

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-1}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-1}} \quad (2.11)$$

dengan: $i = 1, 2, \dots, n$; dan $k = 1, 2, \dots, c$.

7. Memeriksa kondisi berhenti dengan aturan persamaan 2.12.

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

Jika tidak maka $t = t+1$ (2.12)

Ulangi langkah ke-4.

2.7 Indexs Xie dan Beni

Tujuan melakukan proses *clustering* adalah mengumpulkan objek-objek yang memiliki kemiripan yang tinggi dalam satu kluster yang sama. Menurut Xie dan Beni ukuran kevalidan kluster merupakan proses evaluasi hasil klusterisasi untuk menentukan kualitas kluster. Kevalidan suatu kluster merupakan hasil rasio dari kepadatan (*compactness*) dengan keterpisahan (*separation*). Kepadatan adalah ukuran kedekatan antar anggota pada tiap kluster sedangkan keterpisahan adalah ukuran keterpisahan antarkluster satu dengan kluster lainnya. Rasio dari kepadatan dan keterpisahan tersebut didefinisikan sebagai pada persamaan 2.13 , 2.14 dan 2.15.

$$S = \frac{\pi}{N \cdot Dmin} \quad (2.13)$$

$$\pi(\text{Compactness}) = \frac{\sum_{i=1}^n \sum_{k=1}^c \mu_{ik}^w \|x_i - c_j\|^2}{N} \quad (2.14)$$

$$Dmin(\text{Separation}) = \min_{ij} \|c_i - c_j\|^2 \quad (2.15)$$

Dimana $\|x_i - c_j\|$ merupakan jarak *euclid* dari pusat kluster c_j ke poin data x_i dan $\|c_i - c_j\|$ adalah jarak *euclid* dari pusat kluster c ke c , sementara μ adalah derajat

keanggotaan poin data ke-j pada klaster ke-i dan N adalah jumlah data. Semakin kecil nilai S maka semakin bagus hasil klaster yang telah dilakukan [TAN-05].

