

BAB 2

TINJAUAN PUSTAKA

2.1 Tinjauan Non Statistik

2.1.1 Provinsi Jawa Tengah

Jawa Tengah adalah salah satu provinsi di Jawa yang diapit oleh dua provinsi besar yakni Jawa Barat dan Jawa Timur. Jawa Tengah memiliki luas wilayah 3,25 juta hektar atau sekitar 25,04 persen dari luas Pulau Jawa. Secara geografis, Jawa Tengah memiliki batas-batas: Utara-Laut Jawa; Selatan-Provinsi Daerah Istimewa Yogyakarta dan Samudra Hindia; Barat-Provinsi Jawa Barat; Timur-Provinsi Jawa Timur. Jawa Tengah terdiri dari 35 Kabupaten/Kota yaitu 29 Kabupaten dan 6 Kota (BPS Jateng, 2017b).



Gambar 2.1 Peta Provinsi Jawa Tengah

2.1.2 Permasalahan Stunting

Stunting didefinisikan sebagai bentuk kegagalan pertumbuhan yang menyebabkan penundaan baik fisik dan kognitif pada pertumbuhan dan perkembangan anak. *Stunting* juga dikenal sebagai kegagalan pertumbuhan linear, didefinisikan sebagai ketidakmampuan untuk mencapai potensi tinggi badan untuk usia tertentu. Anak-anak yang mengalami *stunting* dalam kehidupannya akan ditantang untuk mencapai ketinggian dan kemampuan kognitif yang sama dengan anak-anak yang tidak mengalami *stunting*, yaitu anak-anak yang pertumbuhan dan perkembangannya optimal. Anak-anak yang mengalami *stunting* memiliki risiko yang lebih tinggi akibat dari kesehatan yang buruk sepanjang hidupnya (Reinhardt dan Fanzo, 2014).

World Health Organization (WHO) mengukur tingkat perkembangan dan status gizi anak menggunakan titik *cutoff Z-score*, yaitu ukuran jarak antara nilai anak dan nilai ekspektasi dari populasi referensi. Seorang balita dikatakan mengalami *stunting* jika hasil pengukuran nilai *Z-Score* kurang dari -2.0 SD (Standar Deviasi). Berdasarkan Keputusan Menteri Kesehatan Nomor 1995/MENKES/XII/2010 tentang Standar Antropometri Penilaian Status Gizi Anak. Tabel klasifikasi status gizi balita berdasarkan nilai *Z-Score* dapat dilihat pada tabel 2.1 dibawah ini:

Tabel 2.1 Klasifikasi Status Gizi Balita Berdasarkan Nilai Z-Score Indeks

Indeks	Kategori Status Gizi	Ambang Batas (Z-Score)
TB/U (<i>Stunting</i>)	Sangat Pendek	Z-Score < -3
	Pendek	-3 < Z-Score < -2
	Normal	-2 < Z-Score < 2
	Tinggi	Z-Score > 2

BB/U (Underweight)	Gizi Buruk	Z-Score < -3
	Gizi Kurang	-3 < Z-Score < -2
	Gizi Baik	-2 < Z-Score < 2
	Gizi Lebih	Z-Score > 2
BB/TB (Wasting)	Sangat Kurus	Z-Score < -3
	Kurus	-3 < Z-Score < -2
	Normal	-2 < Z-Score < 2
	Gemuk	Z-Score > 2

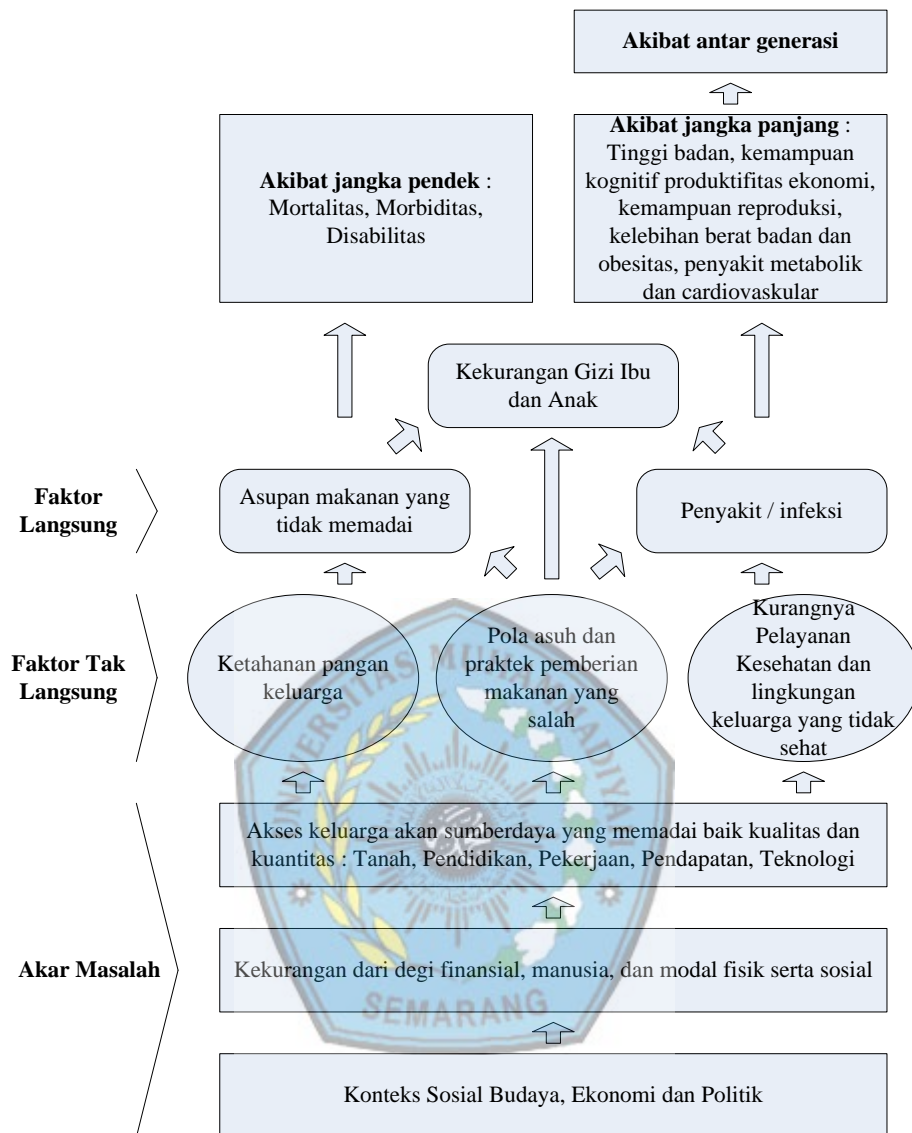
Sumber: Kementerian Kesehatan RI, 2010

Untuk membandingkan permasalahan gizi antar wilayah digunakan angka prevalensi *stunting*, yaitu persentase jumlah balita yang mengalami kejadian *stunting* pada periode tertentu terhadap keseluruhan populasi balita pada wilayah tertentu dan periode tertentu. Menurut WHO klasifikasi permasalahan gizi dapat dilihat pada tabel 2.2 dibawah ini:

Tabel 2.2 Klasifikasi Masalah Gizi Menurut WHO

Indikator	Tingkat Kekurangan Gizi berdasarkan Prevalensi			
	Rendah	Sedang	Tinggi	Sangat Tinggi
TB/U (<i>Stunting</i>)	< 20	20 – 29	30 – 39	≥ 40
BB/U (Underweight)	< 10	10 – 19	20 – 29	≥ 30
BB/TB (Wasting)	< 5	5 – 9	10 – 14	≥ 15

Sumber: WHO (1995)



Gambar 2.2 Kerangka Konseptual status Gizi

Berdasarkan kerangka konseptual UNICEF pada Gambar 2.2, tahapan penyebab timbulnya kekurangan gizi pada anak dibagi menjadi 2 yaitu penyebab langsung dan tidak langsung. Penyebab langsung pertama yaitu asupan makanan yang tidak memadai baik dalam segi kualitas dan kuantitas. Kualitas gizi makanan memegang peranan penting dalam mendorong proses biologi yang mengatur pertumbuhan dan perkembangan syaraf pada anak. Selain itu kualitas makanan merupakan cermin dari keanekaragaman makanan serta kandungan vitamin dan

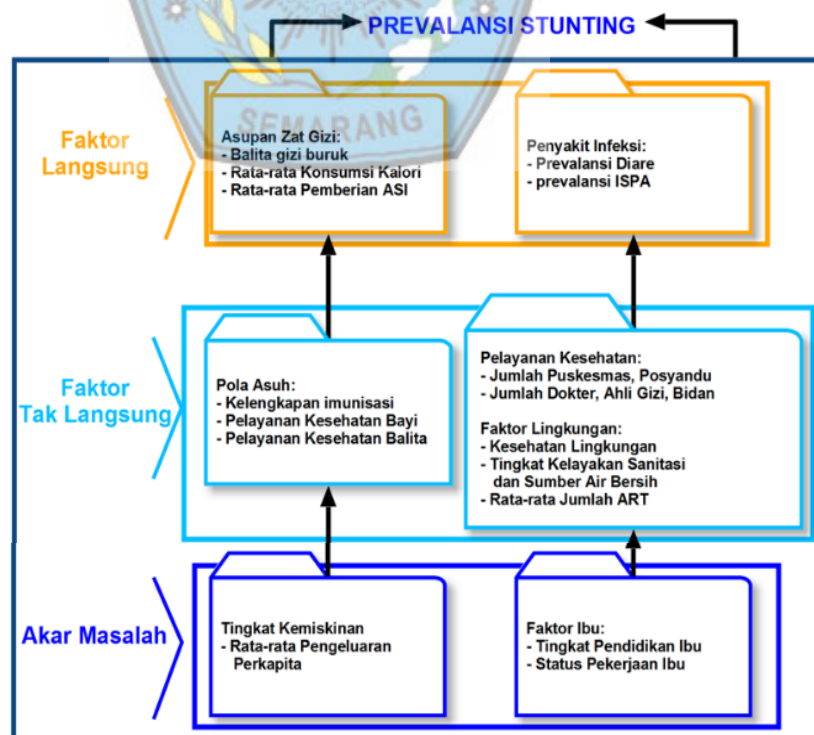
mineral yang dikonsumsi. Penyebab langsung kedua dalam kerangka konseptual adalah penyakit atau infeksi. Penyakit dapat menjadi penyebab dan akibat dari kekurangan gizi. Infeksi masa kanak-kanak dan penyakit diare dapat menyebabkan penyerapan gizi yang buruk atau kemampuan untuk mempertahankan nutrisi yang melemah, dengan begitu akan mengakibatkan resiko *stunting* pada anak-anak. Jika kejadian diare meningkat, resiko *stunting* juga meningkat. Selain diare, penyakit Infeksi saluran Pernapasan Atas (ISPA) juga dapat menyebabkan resiko *stunting* pada anak.

Penyebab tidak langsung kekurangan gizi pada anak berada pada tingkatan masyarakat, dan terdiri dari tiga faktor, penyebab pertama yaitu ketahanan pangan dalam rumah tangga yang tidak memadai. Ketahanan pangan ditandai dengan keberagaman makanan yang tersedia, akses untuk mendapatkan makanan bergizi, dan pemanfaatan makanan yang tepat. Pada tingkat rumah tangga, kerawanan pangan lebih berkaitan dengan ekonomi rumah tangga dan pengetahuan tentang bagaimana makanan digunakan dan dikonsumsi. Beberapa studi tentang efek ketahanan pangan rumah tangga pada anak-anak pra-sekolah menunjukkan bahwa rumah tangga rawan pangan memiliki efek pada prevalensi diare, yang mengarah ke prevalensi *stunting*.

Penyebab tidak langsung kedua adalah pola asuh dan praktik pemberian makanan yang salah, yaitu meliputi kurangnya pemberian ASI eksklusif dan praktik pemberian makanan pada anak-anak yang salah. ASI adalah komponen penting untuk asupan bayi dan kunci untuk membantu membangun faktor kekebalan tubuh yang kuat. UNICEF merekomendasikan ASI eksklusif

selama 6 bulan pertama. Selain itu penyebab tidak langsung yang berasal dari pola asuh yaitu kesadaran orangtua untuk memeriksakan kesehatan anaknya secara rutin ke fasilitas pelayanan kesehatan, kegiatan ini dapat meningkatkan kualitas pertumbuhan dan perkembangan balita.

Penyebab tidak langsung ketiga adalah tidak memadainya layanan kesehatan dan lingkungan rumah tangga yang tidak sehat. Hal itu mencakup sulitnya rumah tangga dalam mendapatkan pelayanan kesehatan serta kualitas yang buruk dari layanan kesehatan, fasilitas air, fasilitas sanitasi, praktek standar, dan persiapan makanan yang tidak memadai. Semua itu sangat penting dalam kontribusi menciptakan lingkungan yang bebas dari infeksi penyakit, dan penyakit merupakan penyebab langsung kedua yang telah dijelaskan di atas. Lebih lengkapnya pemilihan variabel *stunting* balita dapat dilihat di Gambar 2.3 berikut ini:



Gambar 2.3 Kerangka Pemilihan Variabel *Stunting*

Berdasarkan uraian mengenai kerangka konseptual status gizi di atas, maka variabel yang akan digunakan dalam penelitian ini adalah sebagai berikut; variabel dari faktor tidak langsung yang digunakan yaitu kelengkapan imunisasi, kecukupan puskesmas, kecukupan posyandu, kecukupan dokter, kecukupan ahli gizi, kecukupan bidan, cakupan pelayanan kesehatan bayi dan balita, kelayakan sanitasi, kelayakan sumber air bersih dan jumlah anggota rumah tangga. Sedangkan variabel dari faktor langsung yang akan digunakan dalam pengelompokan ini adalah tingkat konsumsi kalori rumah tangga balita, rata-rata lama pemberian ASI, keluhan diare dan ISPA pada balita, bayi dengan berat lahir rendah serta persentase balita dengan gizi buruk dan kurang. Sedangkan variabel yang berasal dari akar masalah yang digunakan yaitu rata-rata pendapatan rumah tangga, tingkat pendidikan ibu, dan status pekerjaan ibu.

2.2 Tinjauan Statistik

2.2.1 Analisis Cluster

Analisis *cluster* merupakan suatu teknik statistik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan kesamaan karakteristik yang dimilikinya. Analisis *cluster* adalah teknik yang digunakan untuk menggabungkan observasi ke dalam *cluster* dengan observasi dalam *cluster* yang sama bersifat homogen atau memiliki kesamaan karakteristik, sementara antara *cluster* yang satu dengan yang lainnya harus memiliki perbedaan terhadap karakteristik tertentu, atau dengan kata lain observasi dalam *cluster* yang satu harus berbeda dengan observasi dalam *cluster* lainnya (Sharma, 1996).

2.2.2 Fuzzy C-Means

Logika *fuzzy* berdasarkan pada himpunan *fuzzy* pertama kali diperkenalkan oleh Zadeh (1965). Pada himpunan tegas suatu objek mempunyai dua kemungkinan, yaitu menjadi anggota suatu himpunan (bernilai 1), atau tidak menjadi anggota suatu himpunan (bernilai 0). Pada himpunan *fuzzy*, objek dapat menjadi anggota semua himpunan atau dengan kata lain derajat keanggotaan dari himpunan *fuzzy* tersebut bersifat kontinyu dimana nilainya berada dalam rentang [0,1]. *Fuzzy clustering* merupakan metode dari analisis *cluster* yang menggunakan prinsip dari logika *fuzzy* untuk menghitung nilai keanggotaan dari setiap objek dalam setiap kelompok (Mason dan Jacobson, 2007). *Fuzzy clustering* (*soft clustering*) merupakan metode tanpa pengawasan (*unsupervised*) yang kuat untuk menganalisis data dan untuk membangun model (Babuska, 2001).

Kebanyakan penelitian mengenai AGD dan logika *fuzzy* menggunakan algoritma Fuzzy C-Means (FCM) (Grekousis dan Thomas, 2011). FCM menghubungkan derajat keanggotaan suatu objek dengan jarak objek tersebut pada pusat kelompoknya. Suatu objek akan cenderung menjadi anggota suatu *cluster* dimana objek itu memiliki derajat keanggotaan tertinggi. Metode ini pertama kali diperkenalkan oleh Dunn (1973) dan kemudian dikembangkan oleh Bezdek (1981). Dalam AGD, FCM lebih banyak digunakan dibandingkan *hard clustering*, karena dapat membantu mengurangi kekeliruan ekologis (Mason dan Jacobson, 2007). Tujuan FCM adalah untuk meminimumkan fungsi objektif, dan FCM merupakan metode yang *robust* dalam meminimumkan fungsi objektif

(Klawonn, 2004) (Klawonn dan Hoppner, 2003). Fungsi Objektif dalam FCM didefinisikan sebagai berikut:

$$J_m(\tilde{U}, v) = \sum_{k=1}^n \sum_{i=1}^m (\mu_{ik})^m (d_{ik})^2$$

$$d_{ik} = d(x_k - v_i) = \left[\sum_{j=1}^m (x_{kj} - v_{ij})^2 \right]^{\frac{1}{2}}$$

dengan:

- μ_{ik} = nilai keanggotaan dari data ke-k pada kelompok ke-i, $0 \leq \mu_{ik} \leq 1$
- d_{ik} = jarak dari titik data x_k ke pusat kelompok v_i
- v_i = nilai pusat kelompok ke-i
- x_k = titik data
- n = jumlah objek penelitian
- c = jumlah kelompok yang diinginkan
- m = *fuzziness*, parameter yang digunakan untuk mengukur tingkat kesamaran dari hasil pengelompokan, $m > 1$

2.2.3 Cluster Validity Index (CVI)

Cluster Validity Index (CVI) adalah sebuah ukuran validitas untuk menemukan jumlah kluster optimal yang sepenuhnya dapat menjelaskan struktur data (K. L. Wu & Yang, 2005). Secara umum ada dua fungsi CVI. Yang pertama CVI dapat digunakan untuk menentukan jumlah kluster, dan yang kedua adalah CVI dapat menjelaskan partisi terbaik yang dilakukan. Ada dua macam pendekatan CVI yaitu tradisional dan modern. Pendekatan tradisional dilakukan dengan menjalankan algoritma klustering secara berulang dengan jumlah kluster

berbeda-beda dan kemudian memilih hasil partisi terbaik dari perhitungan yang dilakukan (Das, Abraham, & Konar, 2009), sedangkan pendekatan modern dilakukan dengan menjalankan algoritma klustering yang secara otomatis menentukan klaster optimalnya (Capitaine & Frélicot, 2012). Pada metode tradisional penentuan klaster optimal diperoleh dari hasil beberapa CVI yang menunjukkan klaster optimalnya (K. L. Wu & Yang, 2005). Misalnya, ada tiga CVI yang digunakan untuk mengevaluasi suatu pengelompokan dan hasilnya ada dua CVI yang menunjukkan klaster empat adalah klaster optimal. Maka dapat disimpulkan bahwa klaster empatlah yang merupakan klaster optimal dari pengelompokan tersebut. Ada dua macam teknik validasi klaster yang dikenal yaitu *comparison indices (external evaluation)* dan *validity indices (internal evaluation)* (Cui et al., 2017). *Comparison indices* digunakan untuk membandingkan metode partisi yang satu dengan yang lainnya atau dengan partisi usulan lainnya. Sedangkan *validity indices* berfokus pada evaluasi setiap calon partisi supaya benar-benar terpisah dengan lainnya (C. Wu et al., 2014).

Penentuan jumlah klaster pada suatu data merupakan teori penting karena memiliki peran penting untuk mendapatkan hasil dugaan yang baik pada analisis klaster. Jika penentuan jumlah klaster tidak sesuai maka akan berdampak pada hasil partisi yang tidak optimal sehingga dapat mengakibatkan kesalahan ketika pengambilan keputusan (Wang & Zhang, 2007). CVI memberikan informasi mengenai kualitas dari partisi yang telah dilakukan dan ada atau tidaknya partisi yang lebih baik. Penentuan klaster optimal dalam CVI yaitu dengan mengukur derajat *compactness* atau kekompakan dan *separation* atau separasi struktur data

pada seluruh kluster. *Compatness* adalah ukuran kedekatan elemen dalam kluster. *Separation* adalah ukuran yang menunjukkan perbedaan antara dua kluster, yakni menghitung jarak antar kluster yang satu dengan kluster yang lain, semakin besar nilainya semakin baik kluster yang terbentuk (Kim, Lee, & Lee, 2004) (Wang & Zhang, 2007). Kluster yang baik adalah kluster yang kompak dan terpisah dari kluster yang lain (K. L. Wu & Yang, 2005). Kluster optimal ditentukan ketika nilai terbesar maupun nilai terkecil dari CVI terbentuk, sesuai kriteria dari masing-masing CVI (Wang & Zhang, 2007).

2.2.4 Partition Entropy Index (PEI)

Bezdek mengusulkan validitas dengan menghitung entropi partisi atau partition entropy. Nilai ini mengevaluasi keteracakan data dalam kluster. Nilainya $0 \leq PEI_K \leq \log_2 K$, nilai yang semakin kecil mempunyai arti bahwa kualitas kluster yang didapat semakin baik. Berikut ini formula untuk menghitung PEI.

$$PEI = -\frac{1}{N} \left(\sum_{i=1}^N \sum_{j=1}^K u_{ij} \times \log_2 u_{ij} \right)$$

N merupakan jumlah data dalam set data, K merupakan jumlah kluster, sedangkan u_{ij} menyatakan nilai keanggotaan data ke-i pada kluster ke-j.

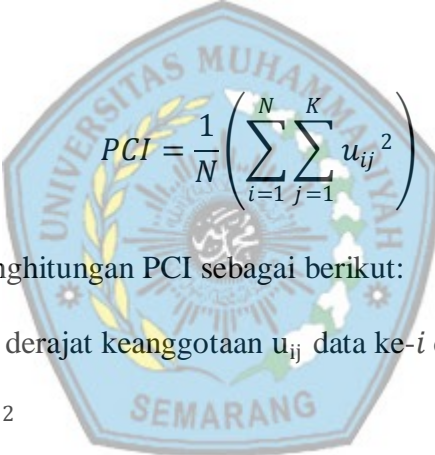
Langkah-langkah penghitungan PEI sebagai berikut:

1. Menghitung nilai derajat keanggotaan u_{ij} data ke-i pada kluster ke-j
2. Mengalikan u_{ij} dengan $\log_2 n$
3. Menghitung $u_{ij} \times \log_2 u_{ij}$
4. Menghitung $\frac{1}{N} \left(\sum_{i=1}^N \sum_{j=1}^K u_{ij} \times \log_2 u_{ij} \right)$

5. Menghitung nilai PEI sesuai persamaan.

2.2.5 Partition Coefficient Index (PCI)

Bezdek (1981) mengusulkan validitas dengan menghitung koefisien partisi atau *partition coefficient* sebagai evaluasi nilai keanggotaan data pada setiap kluster. Nilai ini hanya mengevaluasi nilai derajat keanggotaan, tanpa memandang nilai vektor (data) yang biasanya mengandung informasi geometrik (sebaran data). Nilainya dalam rentang [0,1], nilai yang semakin besar (mendekati 1) mempunyai arti bahwa kualitas kluster yang didapat semakin baik. Berikut formula untuk menghitung PCI:


$$PCI = \frac{1}{N} \left(\sum_{i=1}^N \sum_{j=1}^K u_{ij}^2 \right)$$

Langkah-langkah penghitungan PCI sebagai berikut:

1. Menghitung nilai derajat keanggotaan u_{ij} data ke- i dan kluster ke- j
2. Menghitung $(u_{ij})^2$
3. Menghitung jumlah $(\sum_{i=1}^N \sum_{j=1}^K u_{ij}^2)$
4. Menghitung nilai PCI sesuai persamaan

2.2.6 Fukuyama Sugeno Index (FSI)

Fukuyama dan Sugeno (1989) mengusulkan indeks validitas yang secara bersamaan menghitung derajat keanggotaan fuzzy dan struktur data. Secara umum kluster optimal akan terbentuk ketika nilai FSI semakin kecil yang artinya kluster yang diperoleh semakin baik (Prasetyo, 2014). Penghitungan FSI dilakukan dengan formula sebagai berikut:

$$FSI = \sum_{i=1}^N \sum_{j=1}^K u_{ij}^m \times d(x_i, c_j)^2 - \sum_{i=1}^N \sum_{j=1}^K u_{ij}^m \times d(c_j, \bar{x})^2$$

$$FSI = J_m(u, c) - K_m(u, c)$$

Langkah-langkah penghitungan FSI sebagai berikut:

1. Menghitung nilai derajat keanggotaan u_{ij} data ke- i dan kluster ke- j
2. Menghitung nilai *centroid* masing-masing kluster
3. Menghitung u_{ij}^m , m = bobot pangkat
4. Menghitung jarak data dalam kluster ke- i ke *centroid*-nya $d(x_j, c_i)$ kemudian dikalikan pangkat dua $d(x_j, c_i)^2$
5. Menghitung nilai $J_m(u, c)$ yakni nilai di langkah 3 dikalikan nilai di langkah 4
6. Menghitung nilai FSI sesuai persamaan.

2.2.7 Xie Beni Index (XBI)

Xie dan Beni juga mengusulkan indeks validitas untuk mengevaluasi kluster yang diperoleh namun telah dimodifikasi oleh (Pal & Bezdek, 1995). Secara umum kluster optimal diperoleh ketika nilai XBI semakin kecil yang artinya kluster yang diperoleh semakin baik. Formula yang digunakan sebagai berikut:

$$XBI = \frac{\sum_{j=1}^K \sum_{i=1}^N u_{ij}^m \times d(x_i, c_j)}{N \times \min_{i,j} (d(c_i, c_j)^2)} = \frac{J_m(u, c)/N}{Sep(c)}$$

Langkah-langkah penghitungan XBI sebagai berikut:

1. Menghitung nilai derajat keanggotaan u_{ij} data ke- i dan kluster ke- j
2. Menghitung nilai *centroid* masing-masing kluster
3. Menghitung u_{ij}^m , m = bobot pangkat

4. Menghitung jarak data dalam kluster ke- i ke centroid-nya $d(x_j, c_i)$ kemudian dikalikan pangkat dua $d(x_j, c_i)^2$
5. Menghitung nilai $J_m(u, c)$ yakni nilai di langkah 3 dikalikan nilai di langkah 4
6. Menghitung nilai separasi yakni nilai terkecil dari kuadrat *centroid* antara kluster ke- i dan kluster ke- j kemudian dikalikan $N =$ jumlah data
7. Menghitung nilai XBI sesuai persamaan.

2.2.8 Modification Partition Coefficient Index (MPCI)

Kedua matriks PCI dan PEI mempengaruhi kecenderungan monotonik (berubah secara monoton) terhadap jumlah kluster. Modifikasi dari PCI indeks diusulkan oleh Rajesh N. Dave yaitu *Modification Partition Coefficient Index* (MPCI) yang dapat mengurangi kecenderungan monotonik tersebut. Hasil nilainya dalam rentang $[0,1]$. Kluster optimal akan terbentuk ketika nilai yang diperoleh semakin kecil mendekati nol. Formula yang digunakan seperti berikut:

$$MPCI = 1 - \frac{K}{K-1} (1 - PCI)$$

Langkah-langkah penghitungan MPCI sebagai berikut:

1. Menghitung nilai PCI sesuai persamaan
2. Menghitung nilai $(1 - PCI)$
3. Menghitung nilai $\frac{K}{K-1}$, $K =$ jumlah kluster
4. Menghitung nilai MPCI sesuai persamaan.

2.2.9 Partition Coefficient And Exponential Separation Index (PCAESI)

Wu dan Yang mengusulkan *Partition Coefficient And Exponential Separation Index* (PCAESI) untuk mengevaluasi kluster dengan menghitung

koefisien partisi dan separasi eksponensial untuk kluster ke- i , didefinisikan sebagai gabungan antara kohesi dan separasi kluster tersebut. Kluster optimal diperoleh ketika nilai PCAESI besar. Ukuran kohesi kluster ke- j relatif terhadap kohesi keseluruhan kluster diukur terhadap u_M , seperti pada persamaan berikut: (K. L. Wu & Yang, 2005)

$$Koh_j = \sum_{i=1}^N \frac{u_{ij}^2}{u_M}$$

Sementara u_M didapatkan dari persamaan berikut:

$$u_M = \min_{1 \leq x_j \leq K} \left\{ \sum_{i=1}^N u_{ij}^2 \right\}$$

Separasi kluster ke- j terhadap kluster lain yang terdekat relatif terhadap semua kluster diukur terhadap β_T , seperti persamaan berikut:

$$Sep_j = \exp \left(\frac{\min_{k \neq j} \{d(c_j, c_k)^2\}}{\beta_T} \right), \text{ untuk } k = 1, \dots, K$$

Untuk β_T dinyatakan oleh persamaan berikut:

$$\beta_T = \frac{1}{K} \sum_{i=1}^K d(c_j, \bar{x})^2$$

Untuk menghitung PCAESI pada cluster ke- j dirumuskan seperti berikut:

$$PCAESI = Koh_j - Sep_j$$

Nilai PCAESI _{j} yang besar berarti kluster ke- j bersifat kohesif (kompak) di dalam dan terpisah dari $(K-1)$ kluster yang lain. Nilai yang kecil atau negatif menunjukkan bahwa kluster ke- j dikenali sebagai kluster yang kurang baik. Nilai

kohesi total semua kluster didapatkan dengan menjumlahkan semua nilai kohesi dari setiap kluster, seperti pada persamaan berikut:

$$Koh = \sum_{j=1}^K Koh_j$$

Sementara separasi total semua kluster didapatkan dengan menjumlahkan semua nilai separasi dari setiap kluster, seperti pada persamaan berikut:

$$Sep = \sum_{j=1}^N Sep_j$$

Validitas total dalam PCAESI didefinisikan oleh persamaan berikut:

$$PCAESI = \sum_{j=1}^K PCAESI = Koh - Sep = \sum_{j=1}^K Koh_j - \sum_{j=1}^N Sep_j$$

Nilai PCAESI yang besar berarti K kluster bersifat kohesif (kompak) dan terpisah satu sama lain. Nilai PCAESI yang kecil berarti ada beberapa K kluster yang tidak kompak atau terpisah dari yang lain. Nilai maksimal PCAESI yaitu K , dapat digunakan untuk mendeteksi struktur kluster data dengan partisi kompak dan terpisah dengan baik. Langkah-langkah penghitungan PCAESI sebagai berikut:

1. Menghitung nilai u_M sesuai persamaan
2. Menghitung nilai kohesi sesuai persamaan
3. Menghitung nilai rata-rata kuadrat jarak antara setiap *centroid* terhadap \bar{x} ,
 $d(c_j, \bar{x})^2$
4. Menghitung nilai β_T sesuai persamaan
5. Menghitung jarak *centroid* antara pasangan dua kluster, $d(c_j, c_k)^2$
6. Menghitung nilai separasi sesuai persamaan.