

PENGELOMPOKAN KECAMATAN DI KABUPATEN BANDUNG BERDASARKAN INDIKATOR KEMISKINAN TAHUN 2015 MENGUNAKAN METODE *LATENT CLASS CLUSTER ANALYSIS (LCCA)*

Zahra Ridha Hafizha, Dewi Rachmatin, Entit Puspita

Departemen Pendidikan Matematika FPMIPA UPI

*Surel : zahraridhahafizha@gmail.com

ABSTRAK. Kemiskinan merupakan hal yang sangat kompleks. Kabupaten Bandung merupakan wilayah yang memiliki jumlah penduduk miskin keempat terbanyak di Jawa Barat. Kemiskinan di Kabupaten Bandung mengalami peningkatan dari tahun 2014 ke tahun 2015. Untuk itu diperlukan ketepatan program bantuan dari pemerintah yang bisa diterapkan di Kabupaten Bandung untuk menangani masalah tersebut. Oleh karena itu, penulis akan melakukan pengelompokan kecamatan di Kabupaten Bandung berdasarkan variabel indikator kemiskinan tahun 2015 untuk mendapatkan kelompok-kelompok (klaster) yang memiliki kemiripan sifat antara anggotanya, sehingga untuk masing-masing klaster yang terbentuk memiliki kriteria tertentu. Pengelompokan akan dilakukan dengan metode *Latent Class Cluster Analysis (LCCA)*. Metode ini menggunakan Algoritma EM dan Metode Newton Raphson sebagai tahapan estimasi parameter. Selanjutnya pemilihan model dilakukan berdasarkan nilai BIC terendah dan pengelompokan objek ke dalam klaster dilakukan menggunakan peluang posterior. Kesimpulan dari hasil pengelompokan ini adalah terbentuknya 4 klaster dengan klaster 1 terdiri dari 11 kecamatan, klaster 2 terdiri dari 10 kecamatan, klaster 3 terdiri dari 8 kecamatan dan klaster 4 terdiri dari 2 kecamatan dengan masing-masing klaster memiliki kriteria tertentu.

Kata Kunci : Kemiskinan, *Latent Class Cluster Analysis (LCCA)*, Algoritma EM, Newton-Raphson, Peluang Posterior.

TITLE. Classification Of Subdistrict Of Bandung Regency Based On Poverty Indicator In 2015 Using Latent Class Cluster Analysis (LCCA) Method

ABSTRACT. Poverty is such a complex problem. Bandung Regency takes the fourth place as an area with biggest poverty population in West Java. Poverty in Bandung Regency has increased since 2014 until 2015. Hence, the government needs a precise implementable aid program to help Bandung Regency to handle this problem. Therefore, writer will classify some subdistricts in Bandung Regency based on poverty indicator in 2015 as variable in order to form the cluster that have similar characteristics between each components, so for each cluster that formed will have a specific criteria. Classification process will be done by Latent Class Cluster Analysis (LCCA) method. This method use EM Algorithm and Newton Raphson Method as parameter estimation process. After that, the model will be choosen based on the lowest BIC value and the classification of object into the cluster which will be done by using posterior probability. The conclusion from the results of classification process is four cluster has been obtained with the first cluster consist of 11 subdistricts, second cluster consist of 10 subdistricts, third cluster consist of 8 subdistricts, and fourth cluster consist of 2 subdistricts and each cluster has specific criteria.

Key Words : Poverty, Latent Class Cluster Analysis (LCCA), EM Algorithm, Newton-Raphson, Posterior Probability.

1. PENDAHULUAN

Menurut Badan Pusat Statistika (2016:1), kemiskinan merupakan suatu persoalan yang mendasar dan menjadi pusat perhatian pemerintah di berbagai negara. Kemiskinan dipandang sebagai ketidakmampuan dari segi ekonomi untuk memenuhi kebutuhan dasar makanan dan non makanan yang diukur dari sisi pengeluaran. Suatu penduduk dikatakan miskin apabila memiliki rata-rata pengeluaran per kapita per bulan di bawah Garis Kemiskinan (BPS Jabar, 2017:1). Oleh karena itu, kemiskinan merupakan hal yang sangat kompleks, penanganan dari permasalahan kemiskinan tidaklah sederhana.

Kabupaten Bandung merupakan wilayah yang memiliki jumlah penduduk miskin keempat terbanyak di Jawa Barat. Berdasarkan data dan informasi kemiskinan kabupaten/kota tahun 2014 dan 2015 yang bersumber dari Badan Pusat Statistik, dapat diketahui bahwa jumlah penduduk miskin di Kabupaten Bandung pada tahun 2014 sebanyak 266.800 orang dan tahun 2015 sebanyak 281.000 orang, sehingga kemiskinan di Kabupaten Bandung mengalami peningkatan.

Untuk menanggulangi masalah kemiskinan, diperlukan program pemerintah yang sesuai. Supaya program-program tersebut tepat sasaran sesuai dengan kebutuhan setiap daerahnya, maka penulis ingin melakukan pengelompokan kecamatan di Kabupaten Bandung berdasarkan variabel indikator kemiskinan tersebut. Hal ini bertujuan untuk mendapatkan kelompok-kelompok (klaster) dari kecamatan di Kabupaten Bandung dengan kriteria tertentu untuk masing-masing kelompoknya berdasarkan indikator kemiskinan yang ada, sehingga program bantuan yang akan diterapkan pemerintah Kabupaten Bandung bisa disesuaikan untuk setiap kelompoknya. Alat yang dapat digunakan untuk mengelompokkan data adalah analisis klaster. Dalam kesempatan ini akan digunakan salah satu metode dalam analisis klaster yaitu metode *Latent Class Cluster Analysis (LCCA)*.

Data yang digunakan berasal dari Basis Data Terpadu dan akan digunakan 9 variabel indikator kemiskinan yaitu jumlah individu yang tidak bersekolah (X1), jumlah individu yang mengalami kecacatan (X2), jumlah individu yang memiliki penyakit kronis (X3), jumlah individu usia 15-59 tahun yang tidak bekerja (X4), jumlah rumah tangga yang tinggal di rumah kontrak/sewa (X5), jumlah rumah tangga dengan sumber air minum tidak terlindungi (X6), jumlah rumah tangga yang memasak dengan bahan bakar briket/arang/kayu (X7), jumlah rumah tangga yang tidak memiliki fasilitas BAB (X8) dan jumlah rumah tangga yang pembuangan akhir tinja di lubang tanah (X9).

2. METODOLOGI

Sebelum ke tahapan pembentukan klaster, dilakukan pengujian terhadap asumsi analisis klaster yaitu bebas pencilan dan bebas multikolinearitas. Selain itu dilihat juga apakah data berdistribusi normal multivariat atau tidak. Setelah itu akan didefinisikan data yang digunakan dalam analisis klaster yaitu dimisalkan terdapat p variabel dan n objek, sehingga x_{ih} merupakan respon variabel ke- i objek ke- h dengan $i=1,2,\dots,p$ dan $h=1,2,\dots,n$.

Langkah awal yang dilakukan untuk membentuk klaster adalah pembentukan model LCCA. Sebelum membentuk model, perlu diketahui jenis variabel indikator pada data yang digunakan, sehingga distribusi variabel tersebut disesuaikan. Untuk variabel indikator kontinu menggunakan distribusi

normal multivariat. Sehingga dalam model LCCA, fungsi kepadatan peluang yang digunakan mengikuti sifat ke-2 distribusi multivariat normal (Gustiana dkk., 2015:4), didefinisikan dengan rumus sebagai berikut:

$$g(\mathbf{x}_h | \mu_{ij}, \sigma_{ij}^2) = \frac{1}{(2\pi)^{\frac{p}{2}} [\prod_{i=1}^p \sigma_{ij}^2]} \exp \left[-\frac{1}{2} \sum_{i=1}^p \frac{(x_{ih} - \mu_{ij})^2}{\sigma_{ij}^2} \right] \quad (1)$$

dengan

μ_{ij} = parameter lokasi variabel kontinu X_i di kelas j

σ_{ij}^2 = varians dari variabel ke i di kelas j

j = banyaknya kelas (1,2, ..., c).

Setiap kelas laten/klaster ($j=1,2,\dots,c$) memiliki vektor mean yang terdiri dari rata-rata variabel indikator ke- i di kelas j (μ_{ij}) dan juga memiliki matriks diagonal varians-kovarians yang terdiri dari varians dari variabel indikator ke- i di kelas j .

Kemudian persamaan (1) disubstitusikan ke dalam persamaan model LCCA (2) yang menggunakan distribusi *mixture*. Model LCCA didefinisikan dengan rumus (Vermunt dan Magidson, 2002:2)

$$f(\mathbf{x}_h | \theta) = \sum_{j=1}^c \eta_j g(\mathbf{x}_h | \theta) \quad (2)$$

di mana

j = banyak kelas (1,2, ..., c)

η_j = peluang prior (awal) kelas ke j

$g(\mathbf{x}_h | \theta)$ = fungsi gabungan berupa fungsi kepadatan peluang

dari setiap variabel yang digunakan

θ = himpunan parameter

Sehingga diperoleh rumus sebagai berikut :

$$f(\mathbf{x}_h | \eta_j, \mu_{ij}, \sigma_{ij}^2) = \sum_{j=1}^c \eta_j \left[\frac{1}{(2\pi)^{\frac{p}{2}} [\prod_{i=1}^p \sigma_{ij}^2]} \exp \left[-\frac{1}{2} \sum_{i=1}^p \frac{(x_{ih} - \mu_{ij})^2}{\sigma_{ij}^2} \right] \right] \quad (3)$$

Estimasi parameter dari model LCCA dilakukan dengan metode maksimum likelihood dengan memaksimalkan fungsi log-likelihood. Karena terdapat data yang tidak teramati, digunakan metode iteratif untuk mengestimasi parameter dalam model awal yaitu algoritma *Expectation-Maximization (EM)* dan dilanjutkan dengan menggunakan metode Newton-Raphson (NR). Alasan pemilihan metode tersebut karena kestabilan algoritma EM meskipun nilai taksiran jauh dari nilai optimum dan kecepatan metode NR saat nilai taksiran dekat dengan nilai optimum (Vermunt dan Magidson, 2002: 6). Berikut ini adalah persamaan fungsi log-likelihood :

$$\log L(\theta) = \sum_{h=1}^n \log \left[\sum_{j=1}^c \eta_j \left[\frac{1}{(2\pi)^{\frac{p}{2}} [\prod_{i=1}^p \sigma_{ij}^2]} \exp \left[-\frac{1}{2} \sum_{i=1}^p \frac{(x_{ih} - \mu_{ij})^2}{\sigma_{ij}^2} \right] \right] \right] \quad (4)$$

Tahapan estimasi parameter dilakukan dengan algoritma EM yang memiliki dua tahapan yaitu *E-step* dan *M-step*.

1. *E-Step* digunakan untuk mencari suatu fungsi yaitu ekspektasi dari fungsi log likelihood berdasarkan data terobservasi. Tahapan *E-step* dilakukan untuk mengisi data yang tidak teramati berupa keanggotaan dari setiap kluster.

$$Q = E[\log[L(\theta)] | \mathbf{x}_h, \hat{\theta}^{(r-1)}] \quad (5)$$

2. *M-step* digunakan untuk mencari nilai taksiran parameter yang didefinisikan pada *E-step* untuk memaksimalkan fungsi log likelihood. Berdasarkan persamaan (5), akan dicari nilai taksiran dari $\eta_j^{(r)}, \mu_{ij}^{(r)}, \sigma_{ij}^{2(r)}$ dengan menurunkan secara parsial terhadap masing-masing parameter.

$$\hat{\eta}_j^{(r)} = \sum_{h=1}^n \frac{P(j | \mathbf{x}_h, \hat{\eta}_j^{(r-1)}, \hat{\mu}_{ij}^{(r-1)}, \hat{\sigma}_{ij}^{2(r-1)})}{n} \quad (6)$$

$$\hat{\mu}_{ij}^{(r)} = \frac{\sum_{h=1}^n x_{ih} P(j | \mathbf{x}_h, \hat{\eta}_j^{(r-1)}, \hat{\mu}_{ij}^{(r-1)}, \hat{\sigma}_{ij}^{2(r-1)})}{n \hat{\eta}_j^{(r)}} \quad (7)$$

$$\hat{\sigma}_{ij}^{2(r)} = \frac{\sum_{h=1}^n \sum_{j=1}^c (x_{ih} - \hat{\mu}_{ij}^{(r)})^2 P(j | \mathbf{x}_h, \hat{\eta}_j^{(r-1)}, \hat{\mu}_{ij}^{(r-1)}, \hat{\sigma}_{ij}^{2(r-1)})}{\sum_{h=1}^n \sum_{j=1}^c P(j | \mathbf{x}_h, \hat{\eta}_j^{(r-1)}, \hat{\mu}_{ij}^{(r-1)}, \hat{\sigma}_{ij}^{2(r-1)})} \quad (8)$$

Setelah algoritma EM selesai dilakukan dan diperoleh nilai taksiran parameternya, selanjutnya digunakan metode Newton-Raphson untuk memaksimumkan nilai taksiran parameter tersebut. Metode Newton-Raphson dimulai dengan nilai awal dari parameter $\theta^{(0)}$. Dalam setiap iterasinya, dihitung vektor gradien yang merupakan turunan pertama dari fungsi log-likelihood ($g(\theta)$) dan matriks Hessian (H) yang merupakan matriks turunan kedua. Hasil dari parameter dalam iterasi (r+1) dihitung menggunakan vektor gradien dan invers dari matriks Hessian yang berasal dari iterasi (r) (Snellman, 2008:21).

Setelah melalui tahapan estimasi parameter, diperoleh nilai taksiran parameter $\hat{\eta}_j, \hat{\mu}_{ij}, \hat{\sigma}_{ij}^2$. Nilai dari taksiran parameter tersebut dapat digunakan untuk memaksimumkan nilai fungsi log-likelihood.

Selanjutnya akan dilakukan pemilihan model terbaik dalam model LCCA yang dilakukan dengan salah satu alat yaitu *Bayesian Information Criterion (BIC)*. Rumus dari BIC adalah sebagai berikut (Magidson dan Vermunt, 2004:4) :

$$BIC = -2 \left[\log \left(L(\hat{\theta}) \right) \right] + k \ln(n) \quad (9)$$

Dengan $\log \left(L(\hat{\theta}) \right)$ merupakan nilai fungsi log-likelihood yang sudah maksimum, k merupakan jumlah parameter model dan n merupakan jumlah data pengamatan. Model terbaik adalah model yang memiliki nilai BIC terendah.

Setelah diperoleh model terbaik, langkah selanjutnya adalah pengelompokan objek ke dalam kluster yang dilakukan dengan menghitung peluang posterior dari masing-masing objek terhadap kluster yang terbentuk. Peluang suatu objek masuk pada suatu kluster dinyatakan sebagai berikut (Vermunt dan Magidson, 2002:7) :

$$P(j | \mathbf{x}_h, \hat{\theta}) = \frac{\hat{\eta}_j g(\mathbf{x}_h | \hat{\theta})}{\sum_{j=1}^c \hat{\eta}_j g(\mathbf{x}_h | \hat{\theta})} \quad (10)$$

Pemilihan anggota dari suatu kluster dilakukan dengan melihat nilai peluang posterior tertinggi dari suatu objek untuk setiap klusternya. Selain itu, dihitung juga *classification error* yang didefinisikan dengan rumus sebagai berikut (Chadidjah, 2016:4) :

$$E = \sum_{j=1}^M \frac{N_i}{N} [1 - \max(P(j|\mathbf{x}_h))] \quad (11)$$

dengan N_i merupakan banyak observasi pada kluster ke- i dan M merupakan banyaknya kluster yang terbentuk. Setelah diperoleh anggota untuk masing-masing kluster, selanjutnya dilakukan interpretasi untuk masing-masing kluster yang terbentuk berdasarkan hasil pengujian yang telah dilakukan.

3. HASIL DAN PEMBAHASAN

Setelah dilakukan pengujian asumsi analisis kluster dan mengetahui data berdistribusi normal multivariat, serta semua asumsi dipenuhi, maka dilakukan tahapan pengelompokan data. Pengolahan data dilakukan dengan *software* XLSTAT. Model LCCA dibentuk melalui tahapan estimasi parameter dengan Algoritma EM dan Metode Newton-Raphson. Selanjutnya akan dilakukan pemilihan model terbaik berdasarkan hasil berikut ini:

Tabel 3. 1 Hasil Model LCCA

Banyaknya Klaster	Log-Likelihood	BIC(LL)	Banyaknya parameter
1	-2488,204	5038,220	18
2	-2354,387	4835,832	37
3	-2306,897	4806,097	56
4	-2273,627	4804,804	75
5	-2258,699	4840,193	94
6	-2239,966	4867,972	113
7	-2207,746	4868,778	132
8	-2199,526	4917,584	151
9	-2192,494	4968,766	170

Dapat dilihat pada tabel 3.1 di atas, model terbaik adalah model yang memiliki nilai BIC terkecil yaitu berada pada model dengan 4 klaster sebesar 4804,804. Pada model tersebut diperoleh nilai dari log-likelihood sebesar -2273,627 dan jumlah parameter yang diestimasi sebanyak 75.

Setelah dipilih model terbaik, akan dilihat nilai *classification error* dengan melihat *classification table* sebagai berikut :

Tabel 3. 2 Classification Table Pada Model 4 Klaster

Probabilistik	Klaster1	Klaster2	Klaster3	Klaster4	Total
Klaster1	10,953	0,100	0,000	0,000	11,052
Klaster2	0,047	9,900	0,000	0,000	9,948
Klaster3	0,000	0,000	8,000	0,000	8,000
Klaster4	0,000	0,000	0,000	2,000	2,000
Total	11,000	10,000	8,000	2,000	31,000

Berdasarkan tabel 3.2 di atas, diperoleh jumlah peluang kecamatan yang diklasifikasi dengan benar adalah 30,853 dari 31. Sehingga diperoleh nilai *classification error* sebesar 0,005 atau 0,5%.

Selanjutnya menentukan anggota pada setiap klaster yang terbentuk. Pemilihan anggota dari suatu klaster dilakukan berdasarkan nilai peluang posterior tertinggi dari suatu objek untuk setiap klasternya. Dari hasil yang diperoleh, dapat diketahui kecamatan mana saja yang termasuk ke dalam

klaster 1, klaster 2, klaster 3 dan klaster 4. Berikut ini adalah hasil pengelompokan kecamatan untuk masing-masing klaster:

1. Klaster 1 terdiri dari Kecamatan Ciwidey, Kecamatan Pasirjambu, Kecamatan Cimaung, Kecamatan Kertasari, Kecamatan Solokan Jeruk, Kecamatan Arjasari, Kecamatan Banjaran, Kecamatan Katapang, Kecamatan Kutawaringin, Kecamatan Margaasih dan Kecamatan Cileunyi.
2. Klaster 2 terdiri dari Kecamatan Rancabali, Kecamatan Nagreg, Kecamatan Canguang, Kecamatan Pameungpeuk, Kecamatan Soreang, Kecamatan Margahayu, Kecamatan Dayeuhkolot, Kecamatan Bojongsoang, Kecamatan Cilengkrang dan Kecamatan Cimencyan.
3. Klaster 3 terdiri dari Kecamatan Ibum, Kecamatan Paseh, Kecamatan Cikancung, Kecamatan Cicalengka, Kecamatan Rancaekek, Kecamatan Majalaya, Kecamatan Ciparay dan Kecamatan Baleendah.
4. Klaster 4 terdiri dari Kecamatan Pangalengan dan Kecamatan Pacet.

Setelah dilakukan pengelompokan kecamatan ke dalam 4 klaster yang terbentuk, selanjutnya dilakukan deskripsi masing-masing klaster. Berikut ini adalah deskripsi untuk masing-masing klaster:

1. Klaster 1 menunjukkan bahwa untuk variabel indikator jumlah individu yang tidak bersekolah (X1), jumlah individu yang mengalami kecacatan (X2), jumlah individu yang memiliki penyakit kronis (X3), jumlah individu usia 15-59 tahun yang tidak bekerja (X4) dan jumlah rumah tangga yang memasak dengan bahan bakar briket/arang/kayu (X7) berada pada posisi 2 tertinggi dari 4 klaster; jumlah rumah tangga yang tinggal di rumah kontrak/sewa (X5), jumlah rumah tangga dengan sumber air minum tidak terlindungi (X6), jumlah rumah tangga yang tidak memiliki fasilitas BAB (X8) dan jumlah rumah tangga yang pembuangan akhir tinja di lubang tanah (X9) berada pada posisi 3 tertinggi atau 2 terendah dari 4 klaster.
2. Klaster 2 menunjukkan bahwa untuk variabel indikator jumlah individu yang tidak bersekolah (X1), jumlah individu yang mengalami kecacatan (X2), jumlah individu yang memiliki penyakit kronis (X3), jumlah individu usia 15-59 tahun yang tidak bekerja (X4), jumlah rumah tangga yang memasak dengan bahan bakar briket/arang/kayu (X7) dan jumlah

rumah tangga yang tidak memiliki fasilitas BAB (X8) berada pada posisi tertinggi dari 4 klaster; jumlah rumah tangga yang pembuangan akhir tinja di lubang tanah (X9) berada pada posisi 2 tertinggi dari 4 klaster; jumlah rumah tangga yang tinggal di rumah kontrak/sewa (X5), jumlah rumah tangga dengan sumber air minum tidak terlindungi (X6) berada pada posisi terendah dari 4 klaster.

3. Klaster 3 menunjukkan bahwa untuk variabel indikator jumlah rumah tangga yang pembuangan akhir tinja di lubang tanah (X9) berada pada posisi tertinggi dari 4 klaster; jumlah rumah tangga yang tinggal di rumah kontrak/sewa (X5), jumlah rumah tangga dengan sumber air minum tidak terlindungi (X6), jumlah rumah tangga yang tidak memiliki fasilitas BAB (X8) berada pada posisi 2 tertinggi dari 4 klaster; jumlah individu yang tidak bersekolah (X1), jumlah individu yang mengalami kecacatan (X2), jumlah individu yang memiliki penyakit kronis (X3), jumlah individu usia 15-59 tahun yang tidak bekerja (X4), jumlah rumah tangga yang memasak dengan bahan bakar briket/arang/kayu (X7) berada pada posisi 3 tertinggi atau 2 terendah dari 4 klaster.
4. Klaster 4 menunjukkan bahwa untuk variabel indikator jumlah rumah tangga yang tinggal di rumah kontrak/sewa (X5), jumlah rumah tangga dengan sumber air minum tidak terlindungi (X6) berada pada posisi tertinggi dari 4 klaster; jumlah individu yang tidak bersekolah (X1), jumlah individu yang mengalami kecacatan (X2), jumlah individu yang memiliki penyakit kronis (X3), jumlah individu usia 15-59 tahun yang tidak bekerja (X4), jumlah rumah tangga yang memasak dengan bahan bakar briket/arang/kayu (X7), jumlah rumah tangga yang tidak memiliki fasilitas BAB (X8) dan jumlah rumah tangga yang pembuangan akhir tinja di lubang tanah (X9) berada posisi terendah dari 4 klaster.

4. KESIMPULAN

Berdasarkan hasil dan pembahasan di atas, diperoleh kesimpulan yaitu klaster 1 terdiri dari 11 kecamatan dengan nilai tertinggi pada klaster 1 berada pada variabel indikator jumlah individu yang mengalami kecacatan. Klaster 2 terdiri dari 10 kecamatan dengan nilai tertinggi berada pada variabel indikator jumlah individu yang tidak bersekolah, jumlah individu yang mengalami kecacatan, jumlah individu yang memiliki penyakit kronis, jumlah individu usia 15-59 tahun yang tidak bekerja, jumlah rumah tangga yang memasak

dengan bahan bakar briket/arang/kayu dan jumlah rumah tangga yang tidak memiliki fasilitas BAB. Klaster 3 terdiri dari 8 kecamatan dengan nilai tertinggi berada pada variabel indikator jumlah rumah tangga yang pembuangan akhir tinja di lubang tanah. Klaster 4 terdiri dari 2 kecamatan dengan nilai tertinggi berada pada variabel indikator jumlah rumah tangga yang tinggal di rumah kontrak/sewa dan jumlah rumah tangga dengan sumber air minum tidak terlindungi.

5. DAFTAR PUSTAKA

- Badan Pusat Statistik. (2016). *Data dan Informasi Kemiskinan Kabupaten/Kota 2015*. Jakarta : Badan Pusat Statistik.
- Badan Pusat Statistik. (2015). *Data dan Informasi Kemiskinan Kabupaten/Kota 2014*. Jakarta : Badan Pusat Statistik.
- Badan Pusat Statistik Provinsi Jawa Barat. (2017). *Perkembangan Tingkat Kemiskinan Jawa Barat Maret 2017*. Bandung : Badan Pusat Statistik.
- Chadidjah, A. (2016). *Latent Class Clustering Dalam Pengelompokan Kelurahan Di DKI Jakarta Berdasarkan Ketelantaran Lansia* : Prosiding Seminar Nasional Pendidikan Matematika 2016. Bandung : Universitas Padjajaran.
- Gustiana, dkk.. (2015). *Perbandingan Fuzzy C-Means Clustering Dengan Latent Class Cluster Analysis*. Bandung : Universitas Padjajaran.
- Magidson J. & Vermunt, J.K. (2004). *Latent Class Models*. In D.Kaplan(Ed.),*The Sage Handbook Of Quantitative Methodology For Social Sciences (Pp. 175-198)*. Thousands Oakes : Sage.
- Snellman, M. (2008). *Case Definition Of Pneumococcal Pneumonia - A Latent Class Analysis Approach*. Helsinki : Departement Of Vaccines.
- Vermunt, J. K., & Magidson, J. (2002). *Latent Class Cluster Analysis*. In J. Hagenaaars, & A. Mccutcheon (Eds.), *Applied Latent Class Analysis (Pp. 89-106)*. Cambridge : Cambridge University Press.