

PENGELOMPOKAN KABUPATEN/KOTA DI KALIMANTAN BARAT BERDASARKAN INDEKS PEMBANGUNAN MANUSIA DENGAN *LATENT CLASS CLUSTER ANALYSIS*

Izza Azzahara, Naomi Nessyana Debataraja, Setyo Wira Rizki

INTISARI

Indeks pembangunan manusia (IPM) mengukur pencapaian penduduk dalam mengakses hasil pembangunan untuk memperoleh pendapatan, kesehatan, pendidikan, dan sebagainya. Pembangunan manusia di Kalimantan Barat terus mengalami peningkatan selama periode 2010 hingga 2019. Pengelompokan Kabupaten/Kota di Kalimantan Barat perlu dilakukan sebagai bahan perencanaan dan evaluasi sasaran program pemerintah. Penelitian ini bertujuan untuk menganalisis pengelompokan Kabupaten/Kota di Kalimantan Barat berdasarkan IPM tahun 2019 dengan menggunakan metode latent class cluster analysis (LCCA). Adapun indikator yang digunakan adalah tingkat partisipasi angkatan kerja (TPAK), tingkat pengangguran terbuka (TPT), persentase penduduk miskin, tingkat kedalaman kemiskinan, tingkat keparahan kemiskinan, angka harapan hidup (AHH), rata-rata lama sekolah (RLS), pertumbuhan ekonomi, harapan lama sekolah (HLS), angka melek huruf (AMH). Metode ini menggunakan algoritma expectation maximization (EM) dan metode Newton Raphson sebagai tahapan estimasi parameter. Selanjutnya pemilihan model dilakukan berdasarkan nilai BIC terkecil dan pengelompokan objek kedalam cluster dilakukan menggunakan peluang posterior. Kesimpulan dari hasil pengelompokan ini adalah terbentuknya dua cluster. Cluster pertama terdiri dari 10 Kabupaten di Kalimantan Barat dengan melihat aspek tingkat partisipasi angkatan kerja, persentase penduduk miskin, tingkat kedalaman kemiskinan, tingkat keparahan kemiskinan dengan rata-rata IPM sebesar 65,83 persen. Sedangkan pada cluster dua terdiri dari empat Kabupaten/Kota di Kalimantan Barat dengan melihat aspek tingkat pengangguran terbuka, angka harapan hidup, rata-rata lama sekolah, pertumbuhan ekonomi, harapan lama sekolah, dan angka melek huruf dengan rata-rata IPM sebesar 71,08 persen.

Kata kunci: Indeks Pembangunan Manusia, *Latent Class Cluster Analysis* (LCCA), *Expectation maximization*, *Newton Raphson*, Peluang Posterior.

PENDAHULUAN

Indeks Pembangunan Manusia (IPM) pertama kali diperkenalkan oleh *United Nation Development Program* (UNDP) pada tahun 1990. IPM mengukur pencapaian penduduk dalam mengakses hasil pembangunan untuk memperoleh pendapatan, kesehatan, pendidikan, dan sebagainya. IPM merupakan indikator yang digunakan untuk mengetahui perkembangan pembangunan dalam jangka panjang. Secara umum, pembangunan manusia di Kalimantan Barat terus meningkatkan selama periode 2010 hingga 2019. IPM Kalimantan Barat meningkat dari 61,97 persen pada tahun 2010 menjadi 67,65 persen pada tahun 2019. Selama periode tersebut, IPM di Kalimantan Barat rata-rata tumbuh sebesar 0,98 persen per tahun. Selama periode 2018-2019, IPM Kalimantan Barat berhasil meningkat sebesar 1,00 persen [1].

Pengelompokan Kabupaten/Kota di Kalimantan Barat perlu dilakukan sebagai bahan perencanaan dan evaluasi sasaran program pemerintah. Tujuannya ialah untuk meningkatkan angka pembangunan manusia berdasarkan indikator-indikator yang membentuk IPM. Oleh sebab itu, digunakan salah satu analisis statistik yaitu analisis *cluster* [2].

Peran *clustering* dalam ilmu pengetahuan semakin berkembang pesat seiring dengan meningkatnya kebutuhan. Dalam berbagai penelitian, seperti dibidang psikologi dan pendidikan, terkadang nilai variabel yang diteliti tidak dapat diukur secara langsung, melainkan harus melalui beberapa indikator. Variabel yang tidak dapat diukur secara langsung ini disebut variabel laten. Satu

diantara metode statistik yang digunakan untuk menganalisis variabel laten yaitu *latent class cluster analysis* [3].

Analisis yang dilakukan pada penelitian ini yaitu pengelompokan Kabupaten/Kota di Kalimantan Barat berdasarkan IPM dengan metode *latent class cluster analysis* (LCCA). Analisis ini digunakan untuk mengelompokkan objek-objek penelitian kedalam kelompok-kelompok berdasarkan kesamaan karakteristik objek tersebut.

Variabel yang digunakan pada penelitian ini bersifat kontinu. Adapun variabel tersebut yaitu tingkat partisipasi angkatan kerja (TPAK), tingkat pengangguran terbuka (TPT), persentase penduduk miskin, indeks kedalaman kemiskinan, indeks keparahan kemiskinan, angka harapan hidup (AHH), rata-rata lama sekolah (RLS), pertumbuhan ekonomi, harapan lama sekolah (HLS), dan angka melek huruf (AMH). Secara umum, *output* penelitian ini berupa hasil pengelompokan yang dapat digunakan sebagai acuan untuk menggambarkan distribusi pengelompokan berdasarkan kondisi profil IPM di Kalimantan Barat. Adapun tujuan dari penelitian ini adalah untuk menganalisis pengelompokan Kabupaten/Kota di Kalimantan Barat berdasarkan indikator Indeks Pembangunan Manusia tahun 2019 menggunakan metode *Latent Class Cluster Analysis*.

Tahapan-tahapan dalam pengelompokan objek dengan metode LCCA dimulai dengan mengetahui model awal dari LCCA, selanjutnya dilakukan tahapan estimasi parameter, kemudian dilanjutkan dengan pemilihan model terbaik berdasarkan nilai BIC terkecil dan selanjutnya pengelompokan objek kedalam *cluster* dilakukan menggunakan peluang posterior dari masing-masing objek terhadap *cluster* yang terbentuk. Kemudian dilakukan interpretasi untuk masing-masing *cluster* yang terbentuk berdasarkan hasil analisis yang telah dilakukan.

ANALISIS MULTIVARIAT

Multivariat merupakan suatu teknik statistika yang digunakan untuk melakukan analisis terhadap suatu kondisi. Sederhananya multivariat ialah teknik analisis yang melibatkan banyak variabel. Teknik ini merupakan pengembangan dari teknik Univariat dan Bivariat.

Secara garis besar, analisis multivariat terbagi menjadi dua yaitu metode dependensi dan metode interdependensi. Metode dependensi yaitu teknik yang digunakan untuk menjelaskan hubungan atau pengaruh satu atau banyak variabel terhadap variabel lainnya. Sedangkan metode interdependensi yaitu teknik yang digunakan untuk melakukan pengelompokan, baik terhadap variabel maupun objek [4].

METODE ESTIMASI MAKSIMUM LIKELIHOOD

Metode Maksimum likelihood merupakan metode estimasi yang penting dalam analisis statistik. Tujuan utama dari estimasi maksimum likelihood adalah untuk memberikan hasil pendugaan yang baik bagi θ , terutama jika sampelnya besar. Prosedur estimasi ini menguji apakah estimasi maksimum yang tidak diketahui dari fungsi *likelihood* suatu sampel nilainya sudah memaksimumkan fungsi *likelihood*.

Misalkan x_1, x_2, \dots, x_n adalah variabel acak yang independen dengan fungsi kepadatan peluang $f(x_h|\theta)$ dengan θ merupakan parameter yang di estimasi dan $h = 1, 2, \dots, n$, maka fungsi kepadatan peluang bersama dari variabel acak x_1, x_2, \dots, x_n disebut sebagai fungsi *likelihood* yang dinyatakan dengan $L(\theta)$ dan didefinisikan sebagai berikut [5]:

$$L(\theta) = f(x_1, x_2, \dots, x_n|\theta)$$

$$L(\theta) = f(x_1|\theta)f(x_2|\theta), \dots, f(x_n|\theta)$$

$$L(\theta) = \prod_{h=1}^n f(x_h|\theta) \quad (1)$$

Kemudian fungsi tersebut di log-kan, Fungsi $\log L(\theta)$ ini disebut sebagai fungsi *log-likelihood*. Sehingga bentuk pesamaannya sebagai berikut:

$$\begin{aligned} \log L(\theta) &= \log \prod_{h=1}^n f(x_h|\theta) \\ \log L(\theta) &= \sum_{i=1}^n \log f(x_h|\theta) \end{aligned} \tag{2}$$

Selanjutnya persamaan (2) diturunkan terhadap parameter θ dan menghasilkan nilai θ yang memaksimalkan $\log L(\theta)$. Maksimum dari $\log L(\theta)$ disebut sebagai taksiran maksimum *likelihood* yang dinotasikan dengan $\hat{\theta}$.

LATENT CLASS CLUSTER ANALYSIS (LCCA)

Metode LCCA merupakan teknik pemodelan statistika yang banyak digunakan dalam penelitian ilmu perilaku dan sosial untuk mengidentifikasi keanggotaan kelompok yang tidak teramati antara objek dengan variabel yang diamati. Metode LCCA merupakan suatu metode dalam analisis *cluster* yang mengasumsikan objek masuk kedalam salah satu dari k kelas laten (*cluster*) dengan jumlah dan ukuran *cluster* tidak diketahui sebelumnya. Untuk menyatakan kedekatan objek ke pusatnya, pada analisis LCCA didasarkan pada probabilitas pengelompokan posterior [6].

Model LCCA merupakan turunan dari model variabel laten yang diklasifikasikan berdasarkan jenis variabel yang digunakan. Untuk variabel kontinu dengan variabel laten kategorik disebut *latent profile analysis* (LPA), hanya sebagian kecil yang menggunakan istilah LPA. Nama yang lebih umum digunakan diantaranya *mixture of normal components*, *mixture model clustering*, *model-base clustering*, *latent discriminant analysis*, dan *latent class clustering analysis* (LCCA). Sehingga istilah LPA adalah sama dengan LCCA yang digunakan dalam penelitian ini [7].

Model Latent Class Cluster Analysis (LCCA)

Didefinisikan terdapat p variabel dan n objek, sehingga x_{ih} adalah variabel ke- i objek ke- h dengan $i = 1, 2, \dots, p$ dan $h = 1, 2, \dots, n$. Data tersebut dapat dituliskan sebagai berikut:

Tahap awal yang dilakukan untuk membentuk *cluster* adalah pembentukan model LCCA. Sebelumnya perlu diketahui jenis variabel pada data yang digunakan, sehingga distribusi variabel tersebut dapat disesuaikan. Untuk variabel kontinu menggunakan distribusi normal multivariat. Sehingga dalam model LCCA didefinisikan dengan rumus sebagai berikut [8]:

$$g(x_h|\mu_{ij}, \sigma_{ij}^2) = \frac{1}{(2\pi)^{\frac{p}{2}} [\prod_{i=1}^p \sigma_i^2]} \exp \left[-\frac{1}{2} \sum_{i=1}^p \frac{(x_{ih} - \mu_{ij})^2}{\sigma_{ij}^2} \right] \tag{3}$$

Setiap kelas laten ($j = 1, 2, \dots, c$) memiliki vektor mean yang terdiri dari rata-rata variabel ke- i dikelas j (μ_{ij}) dan memiliki matriks diagonal varians-kovarians yang terdiri dari varians dari variabel ke- i dikelas j . Selanjutnya Persamaan (4) disubstitusikan kedalam persamaan model LCCA yang menggunakan distribusi *mixture*. Model LCCA didefinisikan dengan rumus:

$$f(x_h|\theta) = \sum_{j=1}^c \eta_j g(x_h|\theta) \tag{4}$$

Kemudian substitusikan Persamaan (3) kedalam Persamaan (4), sehingga model LCCA untuk variabel kontinu didefinisikan dengan rumus sebagai berikut:

$$f(x_h|\eta_j, \mu_{ij}, \sigma_{ij}^2) = \sum_{j=1}^c \eta_j \left[\frac{1}{(2\pi)^{\frac{p}{2}} [\prod_{i=1}^p \sigma_i^2]} \exp \left[-\frac{1}{2} \sum_{i=1}^p \frac{(x_{ih} - \mu_{ij})^2}{\sigma_{ij}^2} \right] \right] \tag{5}$$

dengan, x_h adalah nilai dari objek ke- h , η_j adalah peluang prior (awal) kelas ke- j , μ_{ij} adalah parameter lokasi variabel X_i dikelas j , σ_{ij}^2 adalah variansi dari variabel ke- i di kelas j , c adalah banyaknya *cluster* dan j adalah banyaknya kelas $(1, 2, \dots, c)$.

Estimasi Parameter

Estimasi parameter dari model LCCA dilakukan dengan memaksimalkan fungsi *log-likelihood*. Dalam hal ini, adanya variabel laten X yang tidak teramati atau *missing* mengarahkan pilihan pada teknik analisis numerik (iteratif) yaitu algoritma *expectation maximization* (EM) dan dilanjutkan dengan metode *Newton Raphson*. Pemilihan kedua metode tersebut dilakukan karena algoritma EM yang bersifat stabil meskipun nilainya jauh dari optimum, dan metode *Newton Raphson* yang cepat dalam mengestimasi ketika nilainya sudah mendekati optimum [6]. Berikut adalah persamaan fungsi *log-likelihood*:

$$\log L(\theta) = \sum_{h=1}^n \log \left[\sum_{j=1}^c \eta_j \left[\frac{1}{(2\pi)^{\frac{p}{2} \prod_{i=1}^p \sigma_{ij}^2}} \exp \left[-\frac{1}{2} \sum_{i=1}^p \frac{(x_{ih} - \mu_{ij})^2}{\sigma_{ij}^2} \right] \right] \right] \quad (6)$$

Algoritma EM memiliki dua tahapan estimasi parameter yaitu *E-step* dan *M-step*. Langkah-langkah dalam algoritma EM adalah sebagai berikut [9]:

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

$$Q = E[\log[L(\theta)] | x_h, \hat{\theta}^{(r-1)}] \quad (7)$$

2. *M-step*, digunakan untuk mencari nilai taksiran parameter yang didefinisikan pada *E-step* untuk memaksimalkan fungsi *log-likelihood*. Berdasarkan Persamaan (7), diperoleh nilai taksiran dari $\eta_j^{(r)}, \mu_{ij}^{(r)}, \sigma_{ij}^{2(r)}$ dengan cara mendiferensialkan secara parsial terhadap masing-masing parameter.

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

$$\hat{\mu}_{ij}^{(r)} = \frac{\sum_{h=1}^n x_{ih} P(j | x_h, \hat{\eta}_j^{(r-1)}, \hat{\mu}_{ij}^{(r-1)}, \hat{\sigma}_{ij}^{2(r-1)})}{n \hat{\eta}_j^{(r)}} \quad (9)$$

$$\hat{\sigma}_{ij}^{2(r)} = \frac{\sum_{h=1}^n \sum_{j=1}^c (x_{ih} - \hat{\mu}_{ij}^{(r)})^2 P(j | x_h, \hat{\eta}_j^{(r-1)}, \hat{\mu}_{ij}^{(r-1)}, \hat{\sigma}_{ij}^{2(r-1)})}{\sum_{h=1}^n \sum_{j=1}^c P(j | x_h, \hat{\eta}_j^{(r-1)}, \hat{\mu}_{ij}^{(r-1)}, \hat{\sigma}_{ij}^{2(r-1)})} \quad (10)$$

3. Kemudian hitung nilai peluang posterior $\hat{P}(j | x_h, \hat{\theta}^{(r)})$ yang baru berdasarkan nilai taksiran parameter yang diperoleh dari tahap *M-step*, sehingga dapat dituliskan sebagai berikut:

$$\hat{P}(j | x_h, \hat{\eta}_j^{(r)}, \hat{\mu}_{ij}^{(r)}, \hat{\sigma}_{ij}^{2(r)}) = \frac{\hat{P}(j | x_h, \hat{\eta}_j^{(r)}, \hat{\mu}_{ij}^{(r)}, \hat{\sigma}_{ij}^{2(r)})}{\sum_{j=1}^c \hat{\eta}_j^{(r)}, g(x_h | \hat{\mu}_{ij}^{(r)}, \hat{\sigma}_{ij}^{2(r)})} \quad (11)$$

4. Proses *E-step* dan *M-step* dilakukan secara berulang dengan menggunakan nilai taksiran pada persamaan (8), (9), (10), (11) hingga mencapai nilai estimasi parameter yang konvergen atau $\frac{\theta^{(r+1)} - \theta^r}{\theta^r} \leq \varepsilon$.

dengan r merupakan indeks iterasi, $\theta^{(r)}$ adalah himpunan nilai parameter pada iterasi ke- r , $\theta^{(r-1)}$ adalah himpunan nilai taksiran parameter pada iterasi ke- $r - 1$. Setelah diperoleh nilai taksiran parameter algoritma EM, selanjutnya memaksimalkan nilai taksiran parameter tersebut dengan metode *Newton Raphson*.

Metode *Newton Rapshon* merupakan metode iteratif yang biasa digunakan untuk menghitung estimasi maksimum *likelihood*. Metode ini dimulai dengan nilai awal dari parameter $\theta^{(0)}$. Pada setiap iterasinya, dihitung vektor gradien yang merupakan turunan pertama dari fungsi *log-likelihood* ($g(\theta)$) dan matriks Hessian \mathbf{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) [9].

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*.

Pemilihan Model Terbaik

Pemilihan Model terbaik LCCA dilakukan dengan menggunakan *Bayesian Criteria Information* (BIC). Rumus dari BIC adalah sebagai berikut:

$$BIC = -2 [\log(L(\theta))] + k \ln(n) \tag{12}$$

dengan $\log(L(\theta))$ merupakan nilai fungsi *likelihood* yang sudah maksimum, k adalah banyak parameter yang ditaksir dalam model, dan n merupakan jumlah data pengamatan. Adapun kriteria model terbaik yaitu model dengan nilai BIC terkecil dari beberapa solusi k kelompok yang didapat [9].

Pengelompokan dan Interpretasi Cluster

Setelah diperoleh nilai terbaik, selanjutnya dilakukan pengelompokan objek kedalam *cluster* dengan menghitung peluang posterior dari masing-masing objek terhadap *cluster* yang terbentuk. Peluang suatu objek masuk kedalam suatu *cluster* dinyatakan sebagai berikut [9]:

$$P(j|x_h, \hat{\theta}) = \frac{\hat{\eta}_j g(x_h|\hat{\theta})}{\sum_{j=1}^c \hat{\eta}_j g(x_h|\hat{\theta})} \tag{13}$$

dengan $g(x_h|\hat{\theta})$ adalah fungsi gabungan berupa fungsi kepadatan peluang, $\hat{\eta}_j$ adalah estimasi probabilitas prior kelas, $\hat{\theta}$ adalah parameter estimasi, dan x_h adalah variabel independen untuk pengamatan ke- h .

Pemilihan anggota *cluster* dilakukan dengan melihat nilai peluang posterior tertinggi dari suatu objek untuk setiap *clusternya*. Selain itu dihitung *classification error* yang didefinisikan dengan rumus sebagai berikut [10]:

$$E = \sum_{i=1}^M \frac{N_i}{N} [1 - \max(\hat{h}(j|Y_h))] \tag{14}$$

Dengan N_i adalah banyak unit observasi pada *cluster* ke- i , dan M merupakan banyak *cluster* yang terbentuk.

Setelah diperoleh anggota untuk masing-masing *cluster*, selanjutnya dilakukan interpretasi untuk masing-masing *cluster* yang terbentuk berdasarkan hasil pengujian yang telah dilakukan.

HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini adalah data indikator Indeks Pembangunan Manusia (IPM) Kabupaten/Kota di Kalimantan Barat tahun 2019. Data yang digunakan adalah data sekunder yang diperoleh dari Badan Pusat Statistik.

Pada penelitian ini terdapat 10 indikator IPM yang digunakan untuk analisis data dengan metode LCCA. Variabel tersebut diambil berdasarkan beberapa aspek yaitu dari aspek pekerjaan meliputi TPAK dengan satuan persen dan TPT dengan satuan persen. Untuk aspek kemiskinan meliputi Persentase Penduduk Miskin dengan satuan persen, Tingkat Kedalaman Kemiskinan dengan satuan persen, Tingkat Keparahan Kemiskinan dengan satuan persen. Untuk aspek kesehatan meliputi AHH

dengan satuan tahun. Selanjutnya dari aspek pendidikan meliputi RLS dengan satuan tahun, HLS dengan satuan tahun, dan AMH dengan satuan persen. Kemudian dari aspek ekonomi meliputi Pertumbuhan Ekonomi dengan satuan persen.

Model LCCA dibentuk melalui tahapan estimasi parameter dengan Algoritma EM dan Metode *Newton Raphson*. Selanjutnya dilakukan pemilihan model terbaik berdasarkan hasil berikut ini.

Tabel 1 Hasil model LCCA

Banyaknya <i>cluster</i>	BIC
1	491,5856
2	446,1600
3	476,5690
4	496,6331

Berdasarkan Tabel 1, diketahui model terbaik adalah model yang memiliki nilai BIC terkecil yaitu berada pada *cluster* kedua dengan nilai sebesar 446,1600. Setelah mendapatkan jumlah *cluster* yaitu dua *cluster* selanjutnya membentuk karakteristik dua *cluster*. Pembagian karakteristik *cluster* didasarkan pada profil *cluster* sebagai berikut:

Tabel 2 Profil *cluster*

Variabel	<i>Cluster</i> 1	<i>Cluster</i> 2
TPAK	71,3870	64,0626
TPT	3,1890	6,5949
Persentase penduduk miskin	8,9470	4,9625
Indeks kedalaman kemiskinan	1,2890	0,6950
Indeks keparahan kemiskinan	0,2910	0,1500
AHH	71,4260	71,4950
RLS	6,8170	7,8750
Pertumbuhan ekonomi	4,9770	5,2425
HLS	11,9010	13,4625
AMH	92,4650	94,6774

Berdasarkan Tabel 2, karakteristik pada masing-masing *cluster* dapat diuraikan sebagai berikut:

1. Pada *cluster* pertama nilai rata-rata tingkat partisipasi angkatan kerja sebesar 71,3870, persentase penduduk miskin sebesar 8,9470, indeks kedalaman kemiskinan sebesar 1,2890, indeks keparahan kemiskinan sebesar 0,2910. Nilai rata-rata tersebut lebih tinggi dibandingkan dengan *cluster* dua. Sedangkan nilai rata-rata tingkat pengangguran terbuka sebesar 3,1890, angka harapan hidup sebesar 71,4260, rata-rata lama sekolah sebesar 6,8170, pertumbuhan ekonomi sebesar 4,9770, harapan lama sekolah sebesar 11,9010, angka melek huruf sebesar 92,4650. Nilai rata-rata tersebut lebih rendah dibandingkan dengan *cluster* dua. Berdasarkan uraian, diketahui bahwa pada *cluster* pertama nilai rata-rata tertinggi lebih banyak terdapat pada aspek kemiskinan.

2. Pada *cluster* dua nilai rata-rata tingkat pengangguran terbuka sebesar 6,5949, angka harapan hidup sebesar 71,4950, rata-rata lama sekolah sebesar 7,8750, pertumbuhan ekonomi sebesar 5,2425, harapan lama sekolah sebesar 13,4625, angka melek huruf sebesar 94,6774. Nilai rata-rata tersebut lebih tinggi dibandingkan dengan *cluster* pertama. Sedangkan nilai rata-rata tingkat partisipasi angkatan kerja sebesar 64,0626, persentase penduduk miskin sebesar 4,9625, indeks kedalaman kemiskinan sebesar 0,6950, indeks keparahan kemiskinan sebesar 0,1500. Nilai rata-rata tersebut lebih rendah dibandingkan dengan *cluster* pertama. Berdasarkan uraian, diketahui bahwa pada *cluster* pertama nilai rata-rata tertinggi lebih banyak terdapat pada aspek pendidikan.

Berdasarkan analisis karakteristik anggota *cluster*, maka Kabupaten/Kota untuk masing-masing *cluster* dapat dilihat pada tabel berikut:

Tabel 3 *Classification cluster*

Kabupaten/Kota	<i>Cluster</i>	IPM
Sambas	1	67,02
Bengkayang	1	67,57
Landak	1	65,96
Mempawah	2	65,50
Sanggau	1	65,67
Ketapang	1	67,16
Sintang	1	66,70
Kapuas Hulu	1	65,65
Sekadau	1	64,34
Melawi	1	65,54
Kayong Utara	1	62,66
Kubu Raya	2	67,76
Kota Pontianak	2	79,35
Kota Singkawang	2	71,72

Berdasarkan Tabel 4, dapat disimpulkan bahwa anggota *cluster* sebagai berikut:

1. Untuk *cluster* pertama terdiri dari Kabupaten Sambas, Kabupaten Bengkayang, Kabupaten Landak, Kabupaten Sanggau, Kabupaten Ketapang, Kabupaten Sintang, Kabupaten Kapuas Hulu, Kabupaten Sekadau, Kabupaten Melawi, dan Kabupaten Kayong Utara dengan rata-rata IPM sebesar 65,83 persen.
2. Untuk *cluster* dua terdiri dari Kabupaten Mempawah, Kabupaten Kubu Raya, Kota Pontianak, dan Kota Singkawang dengan rata-rata IPM sebesar 71,08 persen.

PENUTUP

Berdasarkan analisis, diketahui bahwa metode *latent class cluster analysis* (LCCA) dapat digunakan untuk pengelompokan indeks pembangunan manusia (IPM) di Kalimantan Barat. Sehingga berdasarkan penelitian terbentuk sebanyak dua *cluster*. *Cluster* pertama terdiri dari 10 Kabupaten di Kalimantan Barat dengan karakteristik IPM yakni tingkat partisipasi angkatan kerja, persentase penduduk miskin, tingkat kedalaman kemiskinan, tingkat keparahan kemiskinan dengan rata-rata IPM sebesar 65,83 persen. Pada *cluster* pertama nilai rata-rata tertinggi lebih banyak terdapat pada aspek kemiskinan. Sedangkan pada *cluster* dua terdiri dari empat Kabupaten/Kota di Kalimantan Barat dengan karakteristik IPM yakni tingkat pengangguran terbuka, angka harapan hidup, rata-rata lama sekolah, pertumbuhan ekonomi, harapan lama sekolah, dan angka melek huruf. *Cluster* dua memiliki

nilai rata-rata IPM sebesar 71,08 persen. Pada *cluster* pertama nilai rata-rata tertinggi lebih banyak terdapat pada aspek pendidikan.

DAFTAR PUSTAKA

- [1]. Badan Pusat Statistik, *Data dan Informasi Indeks Pembangunan Manusia 2019 Kalimantan Barat*. Badan Pusat Statistik; 2020.
- [2]. Talakua, M.W., Leleury, Z.A., dan Talluta, A.W., Analisis *Cluster* dengan Menggunakan Metode *K-Means* untuk Pengelompokan Kabupaten/Kota di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014, *Jurnal Ilmu Matematika dan Terapan*; 2017.
- [3]. Haeruddin, Tirta, I.M., dan Dewi, Y.S., Analisis Regresi Kelas Laten untuk Data Kategorik dengan Satu Kovariat, *Berkala Saintek*; 2013.
- [4]. Usman, H., dan Sobari, N., Aplikasi Teknik Multivariat untuk Riset Pemasaran, Rajagrafindo Persada, Jakarta; 2013.
- [5]. Purba, S.A., Sutarman, dan Darnius, O., *Maximum Likelihood Based on Newton Raphson, Fisher Scoring and Expectation Maximization Algorithm Application on Accident Data*, International Journal of Advance Research (IJAR); 2018.
- [6]. Vermunt, J.K., dan Magidson, J., *Latent Class Cluster Analysis*, Cambridge University Press; 2002.
- [7]. Vermunt, J.K., dan Magison J., *Latent Class Analysis, in the sage encyclopedia of social sciences research methods, sage*; 2004.
- [8]. Gustiana, Y.Y., Suwanda dan Zulhanif, Perbandingan Fuzzy *C-Means Clustering* dengan *Latent Class Clustering Analysis*; 2015.
- [9]. Hafizha, Z.R., Rachmatin, D., dan Puspita, E., Pengelompokan Kecamatan di Kabupaten Bandung Berdasarkan Indikator Kemiskinan Tahun 2015 Menggunakan Metode *Latent Class Cluster Analysis (LCCA)*, *Jurnal Eureka Matika*; 2018.
- [10] Chadidjah, A., *Latent Class Clustering Analysis* dalam Pengelompokan Kelurahan di DKI Jakarta Berdasarkan Keterlantaran Lansia, *Prosiding Seminar Nasional Pendidikan Matematika*, Surakarta; 2016.

IZZA AZZAHARA : Jurusan Statistika FMIPA UNTAN, Pontianak
 izzastatistik14@student.untan.ac.id

NAOMI NESSYANA DEBATARAJA : Jurusan Matematika FMIPA UNTAN, Pontianak
 naominessyana@math.untan.ac.id

SETYO WIRA RIZKI : Jurusan Matematika FMIPA UNTAN, Pontianak
 setyo.wirarizki@math.untan.ac.id
