JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI e-ISSN: 2614-8811
Published by Departement of Mathematics, Hasanuddin University, Indonesia
https://journal.unhas.ac.id/index.php/jmsk/index
Vol. 19, No. 3, May 2023, pp. 460- 472
DOI: 10.20956/j.v19i3.22299

Multinomial Logistic Regression to Model the Combination of Phdi Status and Hdi Status of District/Cities in Kalimantan Island

Regresi Logistik Multinomial untuk Memodelkan Kombinasi antara Status IPKM dan Status IPM Kabupaten/Kota di Pulau Kalimantan

Yusrian Paliling[1], M. Fathurahman[2], Sri Wahyuningsih[3]

[1] *Laboratorium Statistika Terapan, FMIPA, Universitas Mulawarman*
[2,3] *Program Studi S1 Statistika, FMIPA, Universtias Mulawarman*
Email address: *yusrianhaha@gmail.com[1], fathur@fmipa.unmul.ac.id[2], swahyuningsih@fmipa.unmul.ac.id [3]*

Received: 25 August 2022; Accepted: 20 February 2023; Published: 5 May 2023

Abstract

Multinomial Logistics Regression (MLR) is a regression model developed from the Binary Logistics Regression (BLR) model. The response variable of the RLM model has three or more categories and has a multinomial distribution, with the data scale being nominal. The response variable in this study is a combination of the Public Health Development Index (PHDI) status and the Human Development Index (HDI) status of districts/cities in Kalimantan Island, 2018, divided into four categories with category one as a comparison. The predictor variables used were the number of the public health center, the percentage of poor people, economic growth, the pure junior high school participation rate, and the percentage of the population with a minimum of junior high school education. The MLR parameter model was estimated using the Maximum Likelihood Estimation (MLE) method and Newton-Raphson iteration. The hypothesis testing of the MLR model was used by the Likelihood Ratio Test (LRT) method and the Wald test. The best model selection in this study uses the backward method, and the interpretation of the best MLR model uses the odds ratio value. The results showed that the best MLR model is a model that has three predictor variables. The factors that significantly influenced the combination of PHDI status and the HDI status of districts/cities in Kalimantan Island in 2018 were the percentage of poor people, economic growth, and the percentage of people with the minimum level of education in junior high school.

Keywords: PHDI, HDI, LRT, MLE, Newton-Raphson, MLR, Wald test

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Yusrian Paliling, M. Fathurahman, Sri Wahyuningsih

Abstrak

Regresi Logistik Multinomial (RLM) merupakan model regresi yang dikembangkan dari model Regresi Logistik Biner (RLB). Variabel respon model RLM mempunyai tiga atau lebih kategori dan berdistribusi multinomial dengan skala datanya adalah nominal. Variabel respon dalam penelitian yaitu kombinasi antara status Indeks Pembangunan Kesehatan Masyarakat (IPKM) dan status Indeks Pembangunan Manusia (IPM) kabupaten/kota di pulau Kalimantan tahun 2018 yang terbagi dalam empat kategori dengan kategori satu sebagai pembanding. Variabel prediktor yang digunakan yaitu jumlah puskesmas, persentase penduduk miskin, pertumbuhan ekonomi, angka partisipasi murni SMP, dan persentase penduduk yang berpendidikan minimal SMP. Penaksiran parameter model RLM dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) dan iterasi Newton-Raphson. Pengujian hipotesis parameter model RLM menggunakan metode *Likelihood Ratio Test* (LRT) dan uji Wald. Pemilihan model terbaik dalam penelitian ini menggunakan metode *backward* dan interpretasi model RLM terbaik menggunakan nilai *odds ratio*. Hasil penelitian menunjukkan bahwa model RLM terbaik adalah model yang mempunyai tiga variabel prediktor. Faktor-faktor yang berpengaruh signifikan terhadap kombinasi antara status IPKM dan status IPM kabupaten/kota di pulau Kalimantan tahun 2018 adalah persentase penduduk miskin, pertumbuhan ekonomi, dan persentase penduduk yang berpendidikan minimal SMP.

Kata Kunci: RLM, MLE, Newton-Raphson, LRT, Uji Wald, IPKM, IPM

1. PENDAHULUAN

1.1 Latar Belakang

Regresi merupakan salah satu metode yang dapat digunakan dalam pemodelan statistika. Metode ini digunakan untuk memodelkan hubungan antara variabel respon dan variabel prediktor dan memprediksi variabel respon [9]. Jika dilakukan pemodelan terhadap variabel respon berupa data kategorik (*categorical data*), maka dapat digunakan regresi logistik. Regresi logistik terdiri atas regresi logistik biner, regresi logistik multinomial, dan regresi logistik ordinal. Regresi logistik biner digunakan untuk variabel respon yang memiliki dua kategori dan berdistribusi Bernoulli, sedangkan regresi logistik multinomial dan regresi logistik ordinal digunakan untuk variabel respon yang memiliki tiga atau lebih kategori dan berdistribusi multinomial [6].

Regresi logistik yang dikaji dalam penelitian ini adalah Regresi Logistik Multinomial (RLM). RLM merupakan model regresi logistik yang dapat memodelkan hubungan antara satu variabel respon dan satu atau lebih variabel prediktor, dimana variabel respon mempunyai tiga atau lebih kategori dan berskala nominal [6]. Penelitian yang mengkaji dan mengembangkan model RLM telah banyak dilakukan. Penaksiran parameter model RLM dengan metode *Maximum Likelihood Estimation* (MLE) dan iterasi Newton-Raphson dikaji dan dikembangkan oleh Bohning [4]. Yanagihara dkk. [13] mengembangkan metode *the bias-corrected of Akaike's information criterion* untuk pemilihan model RLM terbaik. Sementara itu, pengujian hipotesis parameter model RLM menggunakan metode *Likelihood Ratio Test* (LRT) dan uji Wald dibahas oleh Hosmer dkk. [6] dan Agresti [7]. Seidu [11] menerapkan model RLM untuk memodelkan faktor-faktor yang memengaruhi *antenatal care*.

Penelitian ini bertujuan memodelkan kombinasi antara status Indeks Pembangunan Kesehatan Masyarakat (IPKM) dan status Indeks Pembangunan Manusia (IPM) kabupaten/kota di Pulau Kalimantan tahun 2018 menggunakan RLM. Penelitian ini seperti yang telah dilakukan oleh Fibriyani dkk. [5], tetapi hanya menggunakan model global (RLM) sedangkan Fibriyani dkk. [5] memodelkannya menggunakan model lokal, yaitu *Geographically Weighted Multinomial Logistic Regression* (GWMLR). Model RLM yang digunakan dalam penelitian ini diperoleh dengan menggunakan metode MLE dan iterasi Newton-Raphson. Pengujian hipotesis signifikansi parameter model RLM yang terdiri atas uji simultan dan uji parsial menggunakan metode LRT dan uji Wald.

Penerapan model RLM dalam penelitian ini mengacu pada Program Pembangunan Kesehatan yang tertuang dalam Undang-Undang Nomor 17 tahun 2007 tentang Rencana Pembangunan Jangka Panjang Nasional (RPJPN) tahun 2005–2025 yang menyatakan bahwa dalam rangka mewujudkan sumber daya manusia yang berkualitas dan berdaya saing, maka kesehatan bersama-sama dengan pendidikan dan peningkatan daya beli masyarakat adalah tiga pilar utama untuk meningkatkan kualitas sumber daya manusia. Komposit dari tiga pilar utama ini disebut dengan Indeks Pembangunan Manusia (IPM). IPM merupakan salah satu alat ukur yang dianggap dapat merefleksikan status pembangunan manusia. Indikator kesehatan dalam IPM adalah Angka Harapan Hidup (AHH). AHH merupakan perkiraan lama hidup rata-rata penduduk dari sejak dilahirkan, dengan asumsi tidak ada perubahan pola mortalitas menurut umur. Namun dalam pembangunan manusia dari sektor kesehatan diharapkan tidak hanya mengupayakan agar penduduk dapat mencapai “umur hidup” yang panjang tetapi juga sehat berkualitas dan tidak bergantung pada orang lain. Oleh karena itu, Badan Penelitian dan Pengembangan Kesehatan (Balitbangkes) Kementerian Kesehatan menyusun suatu indeks yang disebut dengan Indeks Pembangunan Kesehatan Masyarakat (IPKM). IPKM merupakan kumpulan indikator kesehatan yang dapat dengan mudah dan langsung diukur untuk menggambarkan masalah kesehatan [8].

1.2 Regresi Logistik Multinomial

Regresi Logistik Multinomial (RLM) merupakan model regresi yang dibangun dengan distribusi multinomial. Variabel respon model RLM mempunyai tiga kategori atau lebih (polikotomus) dan antar kategori tidak mempunyai urutan atau tingkatan (nominal). Misalkan terdapat p variabel prediktor yaitu $X_1, X_2, \ldots, X_p$ maka model RLM dapat dinyatakan sebagai berikut [1] :

$$\text{logit}\left(P(Y_i = j|\mathbf{x}_i)\right) = \ln\left(\frac{P(Y_i = j|\mathbf{x}_i)}{P(Y_i = J|\mathbf{x}_i)}\right) = \mathbf{x}_i^T \boldsymbol{\beta}_j \qquad (1.1)$$

Model RLM dapat diperoleh dengan melakukan estimasi terhadap parameter modelnya menggunakan metode *Maximum Likelihood Estimation* (MLE). Fungsi *likelihood* yang terbentuk dari model RLM adalah sebagai berikut:

$$\begin{aligned} L(\boldsymbol{\beta}) &= \prod_{i=1}^{n}\left(\pi_1(\mathbf{x}_i)^{y_{i1}}\pi_2(\mathbf{x}_i)^{y_{i2}} \ldots \pi_J(\mathbf{x}_i)^{y_{iJ}}\right) \\ &= \prod_{i=1}^{n}\left(\pi_1(\mathbf{x}_i)^{y_{i1}}\pi_2(\mathbf{x}_i)^{y_{i2}} \ldots \pi_J(\mathbf{x}_i)^{1-y_{i1}-\cdots-y_{i,J-1}}\right) \\ &= \prod_{i=1}^{n}\left(\left(\frac{\pi_1(\mathbf{x}_i)}{\pi_J(\mathbf{x}_i)}\right)^{y_{i1}}\left(\frac{\pi_2(\mathbf{x}_i)}{\pi_J(\mathbf{x}_i)}\right)^{y_{i2}} \ldots \left(\frac{\pi_{J-1}(\mathbf{x}_i)}{\pi_J(\mathbf{x}_i)}\right)^{y_{i,J-1}} \pi_J(\mathbf{x}_i)\right). \end{aligned} \qquad (1.2)$$

Selanjutnya menentukan fungsi *log-likelihood* dengan melakukan transformasi *natural logarithm* terhadap fungsi *likelihood* pada persamaan (1.2), yaitu

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n}\left(y_{i1}\ln\left(\frac{\pi_1(\mathbf{x}_i)}{\pi_J(\mathbf{x}_i)}\right) + \cdots + y_{i,J-1}\ln\left(\frac{\pi_{J-1}(\mathbf{x}_i)}{\pi_J(\mathbf{x}_i)}\right) + \ln \pi_J(\mathbf{x}_i)\right) \qquad (1.3)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{J-1} y_{ij} \ln\left(\frac{\pi_j(\mathbf{x}_i)}{\pi_J(\mathbf{x}_i)}\right) + \sum_{i=1}^{n} \ln \pi_J(\mathbf{x}_i)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{J-1} y_{ij} \ln\left(\exp(\mathbf{x}_i^T\boldsymbol{\beta}_j)\right) + \sum_{i=1}^{n} \ln\left(1 + \sum_{j=1}^{J-1} \exp\left(\mathbf{x}_i^T\boldsymbol{\beta}_j\right)\right)^{-1}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{J-1} y_{ij}\left(\mathbf{x}_i^T\boldsymbol{\beta}_j\right) - \sum_{i=1}^{n} \ln\left(1 + \sum_{j=1}^{J-1} \exp\left(\mathbf{x}_i^T\boldsymbol{\beta}_j\right)\right).$$

Penaksir *Maximum Likelihood* (ML) parameter model RLM dapat diperoleh dengan memaksimumkan fungsi *log-likelihood* pada persamaan (1.3) dengan cara menentukan turunan parsial pertama fungsi *log-likelihood* terhadap parameter yang ditaksir kemudian disamakan dengan nol sebagai berikut.

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_j^T} = \sum_{i=1}^{n}\sum_{j=1}^{J-1} y_{ij}\mathbf{x}_i - \sum_{i=1}^{n}\left(\frac{\mathbf{x}_i \exp\left(\mathbf{x}_i^T\widehat{\boldsymbol{\beta}}_j\right)}{1 + \sum_{j=1}^{J-1}\exp\left(\mathbf{x}_i^T\widehat{\boldsymbol{\beta}}_j\right)}\right) = \mathbf{0}. \tag{1.4}$$

Penaksir ML parameter model RLM yang diperoleh berbentuk implisit Oleh karena itu, untuk mendapatkan penaksir ML parameter membutuhkan pendekatan numerik. Salah satu pendekatan numerik yang dapat digunakan adalah metode iterasi Newton-Raphson [1], dengan formula sebagai berikut :

$$\widehat{\boldsymbol{\beta}}_{(t+1)} = \widehat{\boldsymbol{\beta}}_{(t)} - \left[\mathbf{H}\left(\widehat{\boldsymbol{\beta}}_{(t)}\right)\right]^{-1}\mathbf{g}\left(\widehat{\boldsymbol{\beta}}_{(t)}\right), \tag{1.5}$$

dimana

$\widehat{\boldsymbol{\beta}}_{(t)}$ adalah penaksir ML parameter model RLM pada iterasi ke-t, untuk $t = 0,1,2,\ldots,T$ dan T adalah maksimum banyaknya iterasi.

$\mathbf{g}\left(\widehat{\boldsymbol{\beta}}_{(t)}\right)$ adalah vektor gradien penaksir ML parameter model RLM pada iterasi ke-t, dengan elemen-elemennya adalah seperti pada persamaan (1.4) yang dinyatakan sebagai berikut.

$$\mathbf{g}\left(\widehat{\boldsymbol{\beta}}_{(t)}\right) = \left[\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1} \quad \frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2} \quad \cdots \quad \frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_{J-1}}\right]_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}_{(t)}}$$

$\mathbf{H}\left(\widehat{\boldsymbol{\beta}}_{(t)}\right)$ adalah matriks Hessian penaksir ML parameter model RLM pada iterasi ke-t, dengan elemen-elemennya adalah turunan parsial orde kedua fungsi log-likelihood terhadap masing-masing kombinasi parameter dari model RLM, yaitu

$$\mathbf{H}\left(\widehat{\boldsymbol{\beta}}_{(t)}\right) = \begin{bmatrix} \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1 \boldsymbol{\beta}_1^T} & \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1 \boldsymbol{\beta}_2^T} & \cdots & \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1 \boldsymbol{\beta}_{J-1}^T} \\ \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2 \boldsymbol{\beta}_1^T} & \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2 \boldsymbol{\beta}_2^T} & \cdots & \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2 \boldsymbol{\beta}_{J-1}^T} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_{J-1} \boldsymbol{\beta}_1^T} & \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_{J-1} \boldsymbol{\beta}_2^T} & \cdots & \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_{J-1} \boldsymbol{\beta}_{J-1}^T} \end{bmatrix}_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}_{(t)}}$$

dimana

$$\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_j \boldsymbol{\beta}_j^T} = -\sum_{i=1}^{n} \mathbf{x}_i \left[\left(\frac{\mathbf{x}_i \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}_j\right)}{1+\sum_{j=1}^{J-1} \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}_j\right)}\right)\left(\frac{1}{1+\sum_{j=1}^{J-1} \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}_j\right)}\right)\right] \mathbf{x}_i^T.$$

Penaksir ML parameter model RLM didapatkan ketika iterasi telah mencapai kondisi konvergen yaitu $\left\|\widehat{\boldsymbol{\beta}}_{(t+1)} - \widehat{\boldsymbol{\beta}}_{(t)}\right\| \leq \varepsilon$, dimana ε adalah bilangan riil positif yang sangat kecil.

1.3 Pengujian Hipotesis Parameter

Pengujian hipotesis parameter model RLM meliputi uji simultan dan uji parsial. Uji simultan digunakan untuk mengetahui adanya pengaruh variabel prediktor terhadap variabel respon secara bersama-sama. Statistik uji yang digunakan untuk uji simultan adalah statistik *Wilk's lambda* yang dapat diperoleh dengan metode *Likelihood Ratio Test* (LRT), hipotesis untuk uji simultan adalah [7]:

$$\begin{aligned} &H_0 : \beta_{1j} = \beta_{2j} = \cdots = \beta_{pj} = 0, j = 1, 2, \dots, J-1; \\ &H_1 : \text{Paling sedikit ada satu } \beta_{qj} \neq 0, q = 1, 2, \dots, p. \end{aligned} \tag{1.4}$$

Statistik uji yang digunakan sebagai berikut,

$$G^2 = -2 \ln \Lambda = -2 \ln \left(\frac{L(\widehat{\omega})}{L(\widehat{\Omega})}\right) = 2\left(\ln L(\widehat{\Omega}) - \ln L(\widehat{\omega})\right), \tag{1.5}$$

Daerah penolakan H_0 (daerah kritis) untuk hipotesis pada persamaan (1.4) adalah tolak H_0 jika nilai $G^2 > \chi^2_{(\alpha,v)}$, dimana nilai $\chi^2_{(\alpha,v)}$ dapat diperoleh dari tabel distribusi *chi-square* dengan v adalah derajat bebas, $v = p(J-1)$ dan α adalah tingkat signifikansi. Dalam penelitian ini, digunakan nilai α sebesar 10%.

Uji parsial digunakan untuk mengetahui adanya pengaruh dari masing-masing variabel prediktor terhadap variabel respon secara individu. Hipotesis yang digunakan untuk uji parsial adalah sebagai berikut [7]:

$$\begin{aligned} &H_0 : \beta_{qj} = 0, \\ &H_1 : \beta_{qj} \neq 0, j = 1, 2, \dots, J-1; q = 1, 2, \dots, p. \end{aligned} \tag{1.6}$$

Statistik uji yang digunakan untuk uji parsial adalah statistik Wald,

$$Z = \frac{\hat{\beta}_{qj}}{SE(\hat{\beta}_{qj})} \tag{1.7}$$

Statistik Wald pada persamaan (2.17) secara asimtotik berdistribusi normal standar [9], sehingga daerah kritis untuk hipotesis pada persamaan (2.16) adalah tolak H_0 jika nilai $|Z| > Z_{\alpha/2}$, dimana nilai $Z_{\alpha/2}$ dapat diperoleh dari tabel distribusi normal standar.

1.4 Multikolinieritas

Salah satu asumsi yang harus terpenuhi dalam analisis regresi dengan melibatkan variabel prediktor lebih dari satu adalah tidak terjadi multikolinieritas antar prediktor. Hasil estimasi parameter akan memiliki *error* yang sangat besar jika terjadi multikolinieritas dalam model. Multikolinieritas dapat dideteksi dengan menggunakan nilai *Variance Inflation Factors* (VIF) yang dapat diperoleh dengan formula [13]:

$$VIF_q = \frac{1}{1 - R_q^2}, q = 1,2, \dots, p \tag{1.8}$$

R_q^2 adalah koefisien determinasi model regresi antara X_q dengan variabel prediktor yang lain. Jika nilai VIF_q lebih besar dari 10, maka terdapat multikolinieritas [10].

1.5 Pemilihan Model Terbaik

Metode *backward* merupakan proses yang dapat digunakan untuk memilih model terbaik dengan cara mengelminasi variabel prediktor dari model. Elminasi dimulai dari model lengkap kemudian menghapus satu variabel pada satu waktu [3]. Elminasi variabel dipilih dari yang paling tidak berpengaruh.

1.6 Indeks Pembangunan Kesehatan Masyarakat

Indeks Pembangunan Kesehatan Masyarakat (IPKM) merupakan indikator komposit yang menggambarkan kemajuan pembangunan kesehatan, diformulasikan dari data kesehatan berbasis komunitas yaitu Riset Kesehatan Dasar (Riskesdas), Survei Sosial Ekonomi Nasional (Susenas), dan Survei Potensi Desa (Podes). IPKM pertama kali dikembangkan oleh Balitbangkes pada tahun 2010 dengan menggunakan data survei tahun 2007 dan 2008 yang disebut dengan IPKM 2007 [8].

Tujuan dari pengembangan dan penyusunan IPKM adalah untuk memperkaya informasi indikator kesehatan yang dapat menggambarkan keberhasilan pembangunan kesehatan masyarakat dan sebagai salah satu alat monitor keberhasilan pembangunan kesehatan masyarakat melalui penentuan peringkat provinsi dan kabupaten/kota. IPKM dapat dimanfaatkan sebagai [9]:

1. Pembuatan dasar perencanaan program pembangunan kesehatan di kabupaten/kota.
2. Penyusunan bahan advokasi pemerintah pusat ke pemerintah provinsi maupun kabupaten/kota, agar terpacu memperbaiki peringkat dengan melakukan prioritas program kesehatan beserta sumber dayanya.
3. Salah satu kriteria dan pertimbangan penentuan alokasi dana bantuan kesehatan dari pusat ke provinsi atau kabupaten/kota, dan dari provinsi ke kabupaten/kota.

1.7 Indeks Pembangunan Manusia

Menurut *United Nations Development Programs* (UNDP), Indeks Pembangunan Manusia (IPM) mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. Sebagai ukuran kualitas hidup, IPM dibangun melalui pendekatan tiga dimensi dasar. Dimensi tersebut mencakup umur panjang dan sehat, pengetahuan, dan kehidupan yang layak. Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor. Untuk mengukur dimensi kesehatan, digunakan umur harapan hidup saat lahir. Selanjutnya untuk mengukur dimensi pengetahuan digunakan gabungan indikator harapan lama sekolah dan rata-rata lama sekolah. Adapun untuk mengukur dimensi hidup layak digunakan indikator kemampuan daya beli (*purchasing power parity*) [2]. Capaian pembangunan manusia di suatu wilayah pada waktu tertentu dapat dikelompokkan ke dalam empat kelompok, yaitu sangat tinggi (IPM $\geq$ 80), tinggi (70 $\leq$ IPM $<$ 80), sedang (60 $\leq$ IPM $<$ 70), rendah (IPM $<$ 60) [2]. Pengelompokkan ini bertujuan untuk mengorganisasikan wilayah-wilayah menjadi kelompok-kelompok yang sama dalam hal pembangunan manusia.

2. HASIL DAN PEMBAHASAN

2.1 Deskrispi Data Penelitian

Deskripsi data penelitian dinyatakan dalam statistik deskriptif yang meliputi nilai maksimum, nilai minimum, nilai rata-rata, simpangan baku, dan persentase, Berdasarkan perhitungan menggunakan *software R* di dapatkan hasil yang dapat dilihat pada Tabel 2.1.

Tabel 2.1 Deskripsi Variabel Prediktor

Variabel	Maksimum	Minimum	Rata-rata
X_1	33,00	5,00	18,00
X_2	12,83	2,64	6,27
X_3	7,99	-4,10	5,08
X_4	98,82	68,37	81,19
X_5	81,37	35,58	54,25

Berdasarkan Tabel 2.1, tampak bahwa rata-rata jumlah Puskesmas pada kabupaten/kota di Pulau Kalimantan tahun 2018 adalah sebanyak 18 Puskesmas. Wilayah dengan jumlah Puskesmas paling sedikit adalah Kabupaten Tana Tidung, Kalimantan Utara, Kabupaten Mahakam Ulu, Kalimantan Timur, dan Kabupaten Sukamara, Kalimantan Tengah, yaitu sebanyak 5 Puskesmas sedangkan wilayah dengan jumlah Puskesmas terbanyak adalah Kabupaten Kotabaru, Kalimantan Selatan dengan 33 Puskesmas. Jumlah persentase penduduk miskin tertinggi di pulau Kalimantan tahun 2018 terdapat pada Kabupaten Melawi, Kalimantan Barat sebesar 12,83% dan jumlah persentase penduduk miskin terendah terdapat pada Kota Balikpapan, Kalimantan Timur sebesar 2,64% persen. Rata-rata jumlah persentase penduduk miskin sebesar 6,27%. Rata-rata Pertumbuhan ekonomi kabupaten/kota di Pulau Kalimantan tahun 2018 sebesar 5,08% dengan Pertumbuhan ekonomi tertinggi di Kabupaten Sintang, Kalimantan Barat sebesar 7,99% sedangkan Pertumbuhan ekonomi terendah di Kota Bontang, Kalimantan Timur sebesar -4,10%. APM SMP di Pulau Kalimantan tahun 2018 rata-rata sebesar 81,19 dengan APM SMP tertinggi sebesar 98,82 di Kota Singkawang, Kalimantan Barat dan APM SMP terendah sebesar 68,37 di Kabupaten Banjar, Kalimantan Selatan. Untuk rata-rata persentase penduduk yang berpendidikan minimal SMP di Pulau Kalimantan tahun 2018 sebesar 54,25%, persentase terendah di Kabupaten Sambas, Kalimantan Barat sebesar 35,58% dan persentase tertinggi di Kota Palangkaraya, Kalimantan Tengah sebesar 81,37%.

Tabel 2.2 Deskripsi Variabel Respon

Variabel	Frekuensi	Persentase
$Y = 1$	27	48
$Y = 2$	3	5
$Y = 3$	6	11
$Y = 4$	20	36
Jumlah	56	100

Berdasarkan Tabel 2.2, dapat diketahui bahwa jumlah kabupaten/kota terbanyak adalah dengan status IPKM DBK dan IPM sedang sebanyak 27 kabupaten/kota (48%). Jumlah kabupaten/kota paling sedikit adalah dengan status IPKM DBK dan IPM tinggi sebanyak 3 kabupaten/kota (5%).

2.2 Pendeteksian Multikolinieritas

Pendeteksian multikolinieritas dalam penelitian ini menggunakan nilai VIF yang didapatkan melalui perhitungan *software R* yang hasilnya dapat dilihat pada Tabel 2.3.

Tabel 2.3 Nilai VIF dari Variabel Prediktor

Variabel	Nilai VIF
X_1	1,0223
X_2	1,3426
X_3	1,0370
X_4	1,2287
X_5	1,2502

Berdasarkan Tabel 2.3, terlihat bahwa nilai VIF untuk semua variabel prediktor kurang dari 10. Hal ini menunjukan bahwa tidak terjadi multikolinieritas pada variabel prediktor. Sehingga semua variabel prediktor dapat digunakan untuk model RLM.

2.3 Penaksiran Parameter Model RLM

Model awal dari penaksiran parameter model RLM menggunakan metode MLE dan iterasi Newton-Raphson adalah sebagai berikut.

$$g_2(\boldsymbol{x}) = \beta_{02} + \beta_{12}X_1 + \beta_{22}X_2 + \beta_{32}X_3 + \beta_{42}X_4 + \beta_{52}X_5 \quad (2.1)$$
$$g_3(\boldsymbol{x}) = \beta_{03} + \beta_{13}X_1 + \beta_{23}X_2 + \beta_{33}X_3 + \beta_{43}X_4 + \beta_{53}X_5 \quad (2.2)$$
$$g_4(\boldsymbol{x}) = \beta_{04} + \beta_{14}X_1 + \beta_{24}X_2 + \beta_{34}X_3 + \beta_{44}X_4 + \beta_{54}X_5 \quad (2.3)$$

Pada penelitian ini variabel respon kategori satu yang digunakan sebagai pembanding. Perhitungan parameter model RLM menggunakan *software R*, hasil penaksiran dapat dilihat pada Tabel 2.4.

Tabel 2.4 Hasil Penaksiran dan Pengujian Parameter

Parameter	Penaksiran	*Standard Error*	Z	Keputusan
β_{02}	3,4249	21,2194	0,1614	H_0 gagal ditolak
β_{12}	0,0166	0,1781	0,0931	H_0 ditolak
β_{22}	0,5208	0,7020	0,7419	H_0 ditolak
β_{32}	0,5807	1,1765	0,4936	H_0 gagal ditolak
β_{42}	-0,3582	0,3065	-1,1687	H_0 gagal ditolak
β_{52}	0,2828	0,1561	1,8113*	H_0 gagal ditolak
β_{03}	31,7199	18,3217	1,7313*	H_0 gagal ditolak
β_{13}	-0,2577	0,1814	-1,4209	H_0 gagal ditolak
β_{23}	-2,6683	1,6861	-1,5825	H_0 gagal ditolak
β_{33}	-1,9511	1,5390	-1,2678	H_0 gagal ditolak
β_{43}	-0,0180	0,1457	-0,1236	H_0 gagal ditolak
β_{53}	-0,1007	0,2516	-0,4003	H_0 ditolak
β_{04}	-40,6656	23,4205	-1,7363*	H_0 gagal ditolak
β_{14}	0,0765	0,1445	0,5293	H_0 gagal ditolak
β_{24}	0,6060	0,4863	1,2462	H_0 gagal ditolak
β_{34}	-2,6184	1,3730	-1,9070*	H_0 ditolak
β_{44}	0,1525	0,1501	1,0156	H_0 gagal ditolak
β_{54}	0,6514	0,2567	2,5373*	H_0 ditolak

*****) Signifikan pada $\alpha = 0{,}1$.

Berdasarkan Tabel 2.4, maka diperoleh estimasi model RLM sebagai berikut :

$$\hat{g}_2(\boldsymbol{x}) = 3{,}4249 - 0{,}0166\mathrm{X}_1 + 0{,}5208\mathrm{X}_2 + 0{,}5807\mathrm{X}_3 - 0{,}3582\mathrm{X}_4 + 0{,}2828\mathrm{X}_5 \quad (2.4)$$
$$\hat{g}_3(\boldsymbol{x}) = 31{,}7199 - 0{,}2577\mathrm{X}_1 - 2{,}6683\mathrm{X}_2 - 1{,}9511\mathrm{X}_3 - 0{,}0180\mathrm{X}_4 - 0{,}1007\mathrm{X}_5 \quad (2.5)$$
$$\hat{g}_4(\boldsymbol{x}) = -40{,}6656 + 0{,}0765\mathrm{X}_1 + 0{,}6060\mathrm{X}_2 - 2{,}6184\mathrm{X}_3 + 0{,}1525\mathrm{X}_4 + 0{,}6514\mathrm{X}_5 \quad (2.6)$$

Setelah estimasi model RLM didapatkan, maka selanjutnya dilakukan uji hipotesis untuk mengetahui pengaruh variabel prediktor terhadap variabel respon. Terdapat dua tahapan yang dilakukan dalam pengujian ini, yaitu uji simultan menggunakan metode LRT dan uji parsial menggunakan uji *Wald*. Hipotesis uji simultan adalah :

H_0 : $\beta_{1j} = \beta_{2j} = \beta_{3j} = \beta_{4j} = \beta_{5j} = 0, j = 2{,}3{,}4$
H_1 : $Paling\ sedikit\ ada\ satu\ \beta_{qj} \neq 0, q = 1{,}2{,}3{,}4{,}5.$

Berdasarkan hasil perhitungan menggunakan *software R* pada Lampiran 5 diperoleh nilai G^2 sebesar 89,128 yang lebih besar dari nilai $\chi^2_{(0{,}1;15)}$ sebesar 22,31 sehingga H_0 ditolak. Oleh karena itu disimpulkan paling sedikit terdapat satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon atau jumlah puskesmas, persentase penduduk miskin, pertumbuhan ekonomi, APM SMP, dan persentase penduduk yang berpendidikan minimal SMP secara simultan berpengaruh signifikan terhadap kombinasi antara status IPKM dan status IPM kabupaten/kota di pulau Kalimantan tahun 2018. Pengujian hipotesis selanjutnya adalah uji parsial. Hipotesis yang digunakan adalah :

H_0 : $\beta_{qj} = 0$
H_1 : $\beta_{qj} \neq 0, q = 1{,}2{,}3{,}4{,}5; j = 2{,}3{,}4.$

Taraf signifikan yang digunakan pada uji parsial sebesar $0{,}1$ dengan nilai $Z_{0{,}1/2} = 1{,}645$. Pada Tabel 2.4 terlihat beberapa nilai $|Z| > 1{,}645$ yang menunjukkan variabel tersebut berpengaruh signifikan terhadap variabel respon. Variabel-variabel tersebut adalah persentase penduduk miskin, pertumbuhan ekonomi dan persentase penduduk yang berpendidikan minimal SMP. Berdasarkan hasil uji parsial dapat di ketahui terdapat beberapa variabel prediktor yang tidak berpengaruh terhadap variabel respon sehingga dilakukan pemilihan model terbaik dengan metode *backward*. Proses metode *backward* yaitu memodelkan semua variabel prediktor dengan variabel respon, kemudian mengeluarkan satu per satu variabel prediktor yang paling tidak signifikan. Selanjutnya memodelkan kembali variabel prediktor lainnya dengan variabel respon. Hingga didapatkan model akhir yaitu model dengan semua variabel prediktor yang signifikan. Berdasarkan uji parsial yang telah dilakukan pada model lengkap maka variabel pertama yang akan dikeluarkan dari model adalah X_1 karena memiliki nilai Z yang paling kecil. Selanjutnya dilakukan pemodelan ulang dengan melibatkan variabel prediktor X_2, X_3, X_4, X_5. Hasil pemodelan dapat dilihat pada Tabel 2.5.

Tabel 2.5 Hasil Penaksiran Parameter Model RLM

Parameter	Taksiran	*Standard Error*	Keputusan
β^*_{02}	0,6236	19,4343	H_0 gagal ditolak
β^*_{03}	18,3690	12,3265	H_0 gagal ditolak
β^*_{04}	-40,0879	20,1201	H_0 ditolak
β^*_{22}	0,6319	0,6330	H_0 gagal ditolak
β^*_{23}	-2,6854	1,5655	H_0 ditolak
β^*_{24}	0,5612	0,4483	H_0 gagal ditolak

β_{32}^{*}	0,8253	1,1822	H_0 gagal ditolak

Tabel 2.5 Hasil Penaksiran Parameter Model RLM (Lanjutan)

Parameter	Taksiran	*Standard Error*	Keputusan
β_{33}^{*}	-2,0168	1,2155	H_0 ditolak
β_{34}^{*}	-2,2961	1,0967	H_0 ditolak
β_{42}^{*}	-0,3610	0,3041	H_0 gagal ditolak
β_{43}^{*}	-0,0127	0,1056	H_0 gagal ditolak
β_{44}^{*}	0,1586	0,1344	H_0 gagal ditolak
β_{52}^{*}	0,3031	0,1432	H_0 ditolak
β_{53}^{*}	0,0862	0,1454	H_0 gagal ditolak
β_{54}^{*}	0,6308	0,2136	H_0 ditolak

*) Signifikan pada $\alpha = 0{,}1$.

Berdasarkan Tabel 2.5, terlihat beberapa nilai $|Z| > 1{,}645$ yang menunjukkan variabel tersebut berpengaruh signifikan terhadap variabel respon. Variabel-variabel tersebut adalah persentase penduduk miskin, pertumbuhan ekonomi dan persentase penduduk yang berpendidikan minimal SMP. Berdasarkan hasil tersebut maka masih ada variabel prediktor yang dapat di eleminasi dari model karena tidak berpengaruh secara signifikan terhadap varibel respon yaitu variabel X_4, selanjutnya dilakukan pemodelan dengan melibatkan variabel X_2, X_3, X_5. Hasil dari pemodelan disajikan pada Tabel 2.6.

Tabel 2.6 Hasil Penaksiran Parameter Model RLM

Parameter	Taksiran	*Standard Error*	Keputusan
β_{02}^{**}	-21,5420	9,6162	H_0 ditolak
β_{03}^{**}	18,3639	11,8473	H_0 gagal ditolak
β_{04}^{**}	-20,6609	8,0431	H_0 ditolak
β_{22}^{**}	0,2307	0,4165	H_0 gagal ditolak
β_{23}^{**}	-2,6571	1,2359	H_0 ditolak
β_{24}^{**}	0,3148	0,3849	H_0 gagal ditolak
β_{32}^{**}	0,6405	0,9235	H_0 gagal ditolak
β_{33}^{**}	-2,0295	0,9333	H_0 ditolak
β_{34}^{**}	-1,7378	0,7770	H_0 ditolak
β_{52}^{**}	0,2671	0,1270	H_0 ditolak
β_{53}^{**}	0,0658	0,1310	H_0 gagal ditolak
β_{54}^{**}	0,4956	0,1363	H_0 ditolak

*) Signifikan pada $\alpha = 0{,}1$.

Berdasarkan Tabel 2.6 terlihat beberapa nilai $|Z| > 1{,}645$ yang menunjukkan variabel tersebut berpengaruh signifikan terhadap variabel respon. Variabel-variabel tersebut adalah persentase penduduk miskin, pertumbuhan ekonomi dan persentase penduduk yang berpendidikan minimal SMP. Berdasarkan hasil tersebut maka, tahapan eleminasi dapat di hentikan dengan model terbaik dalam pemodelan kombinasi antara status IPKM dan status IPM kabupaten/kota di Pulau Kalimantan tahun 2018 adalah dengan tiga variabel prediktor yaitu, persentase penduduk miskin, pertumbuhan ekonomi dan persentase penduduk yang berpendidikan minimal SMP.

2.4 Interpretasi Model RLM Terbaik

Interpretasi model RLM terbaik dilakukan mengunakan nilai odds rasio untuk melihat kecenderungan dan juga memperhatikan tanda positif atau negatif pada setiap nilai taksiran dalam model. Nilai odds rasio disajikan pada Tabel 2.7.

Tabel 2.7 Nilai Odds Rasio Model RLM Terbaik

Variabel	**Parameter**	***OR***
Intercept	β_{02}^{**}	$4{,}4100\times10^{-10}$
	β_{03}^{**}	$9{,}4477\times10^{7}$
	β_{04}^{*}	$1{,}0644\times10^{-9}$
Penduduk miskin	β_{22}^{**}	1,2594
	β_{23}^{**}	$7{,}0153\times10^{-2}$
	β_{24}^{**}	1,3700
Pertumbuhan ekonomi	β_{32}^{**}	1,8974
	β_{33}^{**}	0,1314
	β_{34}^{**}	0,1759
Persentase penduduk yang berpendidikan minimal SMP	β_{52}^{**}	1,3062
	β_{53}^{**}	1,0680
	β_{54}^{**}	1,6414

Interpretasi dari model RLM terbaik menggunakan nilai odds rasio pada Tabel 2.7, adalah sebagai berikut :

1. Variabel Penduduk Miskin
   - Setiap peningkatan 1 persen penduduk miskin maka kabupaten/kota memiliki kecenderungan sebesar 1,2594 kali mempunyai status IPKM DBK dan IPM tinggi dari pada IPKM DBK dan IPM sedang.
   - Setiap peningkatan 1 persen penduduk miskin maka kabupaten/kota memiliki kecenderungan sebesar $7{,}0153 \times 10^{-2}$ kali mempunyai status IPKM BDBK dan IPM sedang dari pada IPKM DBK dan IPM sedang.
   - Setiap peningkatan 1 persen penduduk miskin maka kabupaten/kota memiliki kecenderungan sebesar 1,3700 kali mempunyai status IPKM BDBK dan IPM tinggi dari pada IPKM DBK dan IPM sedang.

2. Variabel Persentase Pertumbuhan Ekonomi
   - Setiap peningkatan 1 persen pertumbuhan ekonomi, maka kabupaten/kota memiliki kecenderungan 1,8974 kali mempunyai status IPKM DBK dan IPM tinggi dari pada IPKM DBK dan IPM sedang.
   - Setiap peningkatan 1 persen pertumbuhan ekonomi, maka kabupaten/kota memiliki kecenderungan 0,1314 kali mempunyai status IPKM BDBK dan IPM sedang dari pada IPKM DBK dan IPM sedang.
   - Setiap peningkatan 1 persen pertumbuhan ekonomi, maka kabupaten/kota memilikikecenderungan 0,1759 kali mempunyai status IPKM BDBK dan IPM tinggi dari pada IPKM DBK dan IPM sedang.

3. Variabel persentase penduduk yang berpendidikan minimal SMP

- Setiap peningkatan 1 persen penduduk yang berpendidikan minimal SMP maka kabupaten/kota memiliki kecenderungan lebih besar 1,306 kali mempunyai status IPKM DBK dan IPM tinggi dari pada IPKM DBK dan IPM sedang.
- Setiap peningkatan 1 persen penduduk yang berpendidikan minimal SMP maka kabupaten/kota memiliki kecenderungan lebih besar 1,0680 kali mempunyai status IPKM BDBK dan IPM sedang dari pada IPKM DBK dan IPM sedang.
- Setiap peningkatan 1 persen penduduk yang berpendidikan minimal SMP maka kabupaten/kota memiliki kecenderungan sebesar 1,6414 kali mempunyai status IPKM BDBK dan IPM tinggi dibanding IPKM DBK dan IPM sedang.

3 KESIMPULAN DAN SARAN

3.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan, diperoleh kesimpulan sebagai berikut:

1. Model RLM terbaik pada kombinasi antara status IPKM dan status IPM kabupaten/kota di Pulau Kalimantan tahun 2018 adalah model RLM yang mempunyai tiga variabel prediktor dengan persamaan model sebagai berikut.

$$\hat{g}_2^{**}\ (\boldsymbol{x}) = -21{,}5420 + 0{,}2307X_2 + 0{,}6405X_3 + 0{,}2671X_5$$
$$\hat{g}_3^{**}\ (\boldsymbol{x}) = 18{,}3639 - 2{,}65714X_2 - 2{,}0295X_3 + 0{,}0658X_5$$
$$\hat{g}_4^{**}\ (\boldsymbol{x}) = -20{,}6609 + 0{,}3148X_2 - 1{,}7378X_3 + 0{,}4956X_5$$

2. Faktor-faktor yang berpengaruh signifikan terhadap kombinasi antara status IPKM dan status IPM kabupaten/kota di Pulau Kalimantan tahun 2018 adalah adalah persentase penduduk miskin, pertumbuhan ekonomi dan persentase penduduk yang berpendidikan minimal SMP.

5.2 Saran

Saran dari penelitian ini adalah sebagai berikut:

1. Untuk penelitian selanjutnya pada variabel prediktor dapat menggunakan skala data yang berbeda seperti skala data kategorik.
2. Pendekatan numerik untuk penaksiran parameter model RLM untuk penelitian selanjutnya dapat menggunakan metode *Quasi-Newton* dengan algoritma *Broyden-Fletcher-Goldfarb-Shanno* (BFGS).

DAFTAR PUSTAKA

[1] Agresti, A., 2013. *Categorical Data Analysis*. New York: John Wiley & Sons.

[2] Badan Pusat Statistik.,2019. *Indeks Pembangunan Manusia 2018*. Jakarta: BPS.

[3] Bilder C. R. & Loughin T. M., 2015. *Analysis of Categorical Data with R*. CRC Press, Boca Raton.

[4] Bohning, D., 1992. Multinomial Logistic Regression Algorithm. *Ann. Inst. Statist. Math*, Vol. 44, No. 1, 197 – 200.

[5] Fibriyani V., Latra I. N. & Purhadi., 2015. Geographically Weighted Multinomial Logistic Regression Model (Case Study: Human Development Index Value and Healths Status Areas of Districts/Cities 2013 in Sumatera. *Proceedings of the IconSSE FSM SWCU (pp. 13-19)*. Salatiga, Indonesia: FSM SWCU.

[6] Garson G. D., 2014. *Logistic Regression: Binary & Multinomial.* Statistical Associates Publishing, North Carolina.

[7] Hosmer D. W., Lemeshow S., & Sturdivant R. X., 2013. *Applied Logistic Regression.* John Wiley & Sons Inc., New Jersey.

[8] Kementerian Kesehatan., 2010. *Indeks Pembangunan Kesehatan Masyarakat*. Badan Penelitian dan Pengembangan Kesehatan, Jakarta.
[9] Kementerian Kesehatan., 2019. *Indeks Pembangunan Kesehatan Masyarakat*. Badan Penelitian dan Pengembangan Kesehatan, Jakarta.
[10] Kutner M. H., Nachtsheim C. J., & Neter J., 2004. *Applied Linear Regression Models.* McGraw-Hill/Irwin, New York.
[11] Pawitan Y., 2001. *All Likelihood: Statistical Modelling and Inference Using Likelihood*. Clarendon Press, Oxford.
[12] Seidu A. A., 2021. A Multinomial Regression Analysis of Factors Associated with Antenatal Care Attendance Among Women in Papua New Guinea. *Public Health in Practice*, Vol. 2, No. 100161, 1 – 7.
[13] Suyono., 2015. *Analisis Regresi Untuk Penelitian*. Deepublish, Yogyakarta.
[14] Yanagihara K., Kamo, K., Imory, S., & Satoh, K., 2012. Biased-corrected AIC for Selecting Variables in Multinomial Logistic regression models. *Linear Algebra and Its Applications*, Vol. 436, 4329 – 4341.