

BOOTSTRAP AGGREGATING (BAGGING) REGRESI LOGISTIK ORDINAL UNTUK MENGLASIFIKASIKAN STATUS GIZI BALITA DI KABUPATEN KLUNGKUNG

Palupi Purnama Sari^{1§}, Made Susilawati², I Gusti Ayu Made Srinadi³

¹Jurusan Matematika, Fakultas MIPA – Universitas Udayana [Email: palupi1412@gmail.com]

²Jurusan Matematika, Fakultas MIPA – Universitas Udayana [Email: susilawati.made@gmail.com]

³Jurusan Matematika, Fakultas MIPA – Universitas Udayana [Email: srinadi@unud.ac.id]

[§]Corresponding Author

ABSTRACT

This research was conducted to determine the variables that significantly influence nutritional status of children based on indicators that defined as height for age (H/A) and to classify children nutritional status into normal, short or very short categories. Height for age (H/A) is indicator used to describe the circumstances of malnutrition short. Short children (stunting) is children who fail to reach optimal growth. The secondary data was list of 116 data of children aged 24-59 months at UPT. Puskesmas Klungkung I in 2015. The method was used was ordinal logistic regression and bagging ordinal logistic regression. Based on the research results, it was obtained variables children body length at birth, birth weight, and length of mid-upper arm circumference (MUAC) in pregnant woman were significantly affects the nutritional status of children by the classification accuracy level of ordinal logistic regression 62,86% and misclassification 37,14%. Classification accuracy of ordinal logistic regression can be improved by bagging ordinal logistic regression method. Bagging works well on classification method which has unstable procedures. One of classification method which has unstable procedures is ordinal logistic regression. Bagging ordinal logistic regression method by 501 times replication capable to improve classification accuracy of ordinal logistic regression model from 62,86% to 68,57%, increased 5,71%.

Keywords: Bootstrap Aggregating (bagging), Ordinal Logistic Regression

1. PENDAHULUAN

Klasifikasi regresi logistik yang didasarkan pada pendekatan analisis regresi linear, merupakan metode yang sangat populer dalam mengatasi masalah-masalah klasifikasi (Subianto, 2009). Regresi logistik dengan variabel respons kategori berskala ordinal biasa disebut dengan regresi logistik ordinal. Ketepatan suatu klasifikasi regresi logistik ordinal dapat dihitung dengan menggunakan alat ukur *apparent error rate* (APER) (Johnson & Winchern, 2007). Peningkatan ketepatan klasifikasi dapat dilakukan dengan pendekatan *bootstrap aggregating* (*bagging*). *Bagging* bekerja dengan baik pada metode klasifikasi yang menghasilkan parameter tidak stabil, yang mana perubahan kecil pada *data set* akan

menghasilkan perubahan besar pada model yang diperoleh (Breiman, 1996). Salah satu metode klasifikasi yang tidak stabil adalah regresi logistik ordinal.

Pada penelitian sebelumnya, berkaitan dengan peningkatan klasifikasi dengan metode *bagging* telah dilakukan oleh Insani, et al (2015) yang mengaplikasikan metode *bootstrap aggregating* regresi logistik ordinal untuk memperoleh model dan peningkatan ketepatan klasifikasi regresi logistik ordinal pada kasus kanker serviks. Selanjutnya, Akbar, et al (2010) melakukan perbandingan klasifikasi status gizi balita dengan metode *cluster* dan WHO-NCHS berdasarkan survei kekurangan energi protein (KEP) di Kabupaten Nganjuk, untuk memperoleh model dan peningkatan klasifikasi dilakukan dengan pendekatan *bagging* regresi

logistik ordinal. Berdasarkan penelitian sebelumnya, diperoleh hasil bahwa *bagging* dapat meningkatkan ketepatan klasifikasi dari model *data set* tunggal (regresi logistik ordinal).

Penelitian ini dilakukan untuk mengaplikasikan metode *bootstrap aggeragting (bagging)* regresi logistik ordinal untuk mendapatkan model dan mengklasifikasi status gizi balita di Kabupaten Klungkung ke dalam kategori normal, pendek dan sangat pendek. Kemudian, dibandingkan ketepatan klasifikasi antara regresi logistik ordinal dan *bagging* regresi logistik ordinal untuk dilihat adanya peningkatan tingkat ketepatan klasifikasi.

Salah satu indikator kesehatan yang dinilai pencapaiannya dalam *Millenium Development Goals* (MDG's) adalah status gizi balita (Dinas Kesehatan Provinsi Bali, 2014) Status gizi berdasarkan indikator tinggi badan menurut umur (TB/U) digunakan untuk menggambarkan keadaan kurang kronis, yaitu gizi pendek (*stunting*). Prevalensi balita *stunting* di Provinsi Bali meningkat sebesar 1,6% yaitu dari 31% pada tahun 2007 menjadi 32,6% pada tahun 2013 di Indonesia. Hasil Riskesdas tahun 2013, Kabupaten Klungkung memiliki prevalensi *stunting* sebesar 19,3% yang berada pada urutan kelima di Provinsi Bali (Kementerian Kesehatan, 2013).

Analisis data kategori menggunakan regresi logistik bertujuan untuk mendapatkan model terbaik dan sederhana untuk menjelaskan hubungan antara keluaran dari variabel respons (Y) dengan variabel-variabel prediktornya (X) (Hosmer & Lemeshow, 2000). Regresi Logistik adalah model regresi yang digunakan apabila variabel respons bersifat kualitatif. Suatu variabel respons Y berskala ordinal dapat terdiri dari $K + 1$ nilai dinyatakan dengan $0, 1, 2, \dots, K$. Ekspresi umum peluang bersyarat $Y = k$ pada vektor x dari p variabel prediktor adalah $P[Y = k|x] = \phi_k(x)$. Misalkan $\phi_k(x) = \pi_k(x)$, maka untuk $K = 0, 1, 2$, model yang terbentuk dijelaskan oleh persamaan berikut:

$$P(Y = 0|x) = \frac{1}{1 + e^{g_1(x)} + e^{g_2(x)}} = \phi_0(x)$$

$$P(Y = 1|x) = \frac{e^{g_1(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}} = \phi_1(x)$$

$$P(Y = 2|x) = \frac{e^{g_2(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}} = \phi_2(x),$$

sehingga bentuk persamaan tersebut adalah sebagai berikut:

$$P(Y = k|x) = \frac{e^{g_k(x)}}{\sum_{i=1}^K e^{g_i(x)}}, \quad (1)$$

dengan $g_0(x) = 0$, dan $k = 0, 1, \dots, K$.

Pada regresi logistik ordinal, model dapat diperoleh dengan model odds proporsional (*proportional odds model*). Model logit ini merupakan model yang didapatkan dengan membandingkan peluang kumulatif yaitu peluang kurang dari atau sama dengan kategori respons ke- k pada p variabel prediktor yang dinyatakan dalam vektor x , $P(Y \leq k|x)$, dengan peluang yang lebih besar dari kategori respons ke- k , $P(Y > k|x)$ yang didefinisikan sebagai berikut:

$$\begin{aligned} \text{logit } P(Y \leq k|x) &= c_k(x) = \ln \left[\frac{P(Y \leq k|x)}{P(Y > k|x)} \right] \\ &= \ln \left[\frac{\phi_0(x) + \phi_1(x) + \dots + \phi_k(x)}{\phi_{k+1}(x) + \phi_{k+2}(x) + \dots + \phi_K(x)} \right] \\ &= \tau_k + x'\beta, \end{aligned} \quad (2)$$

untuk $k = 0, 1, 2, \dots, K - 1$.

Metode yang paling umum digunakan untuk menduga parameter pada model regresi logistik adalah metode maksimum likelihood (*method of maximum likelihood*). Bentuk umum dari likelihood untuk sampel dari n amatan yang saling bebas $(z_i, x_i), i = 1, 2, \dots, n$; adalah sebagai berikut

$$l(\beta) = \prod_{i=1}^n [\phi_0(x_i)^{z_{0i}} \times \phi_1(x_i)^{z_{1i}} \times \dots \times \phi_K(x_i)^{z_{Ki}}], \quad (3)$$

dengan $\phi_k(x)$ merupakan fungsi dari parameter yang tidak diketahui dan $z' = (z_0, z_1, \dots, z_k)$ dibentuk dari respons yang ordinal.

Metode kemungkinan maksimum memberikan nilai penduga dari vektor $\beta_k' = (\beta_{0k}, \beta_{1k}, \dots, \beta_{pk})$ dengan memaksimumkan fungsi kemungkinan bersama pada persamaan (3). Logaritma dari fungsi kemungkinan bersamanya dapat ditulis sebagai berikut:

$$L(\beta) = \sum_{i=0}^n z_{0i} \ln [\phi_0(x_i)] + \dots + z_{Ki} \ln [\phi_{Ki}(x_i)].$$

Pengujian terhadap parameter-parameter estimasi model dilakukan untuk mengetahui peran seluruh variabel prediktor baik secara

simultan (bersama-sama) maupun secara parsial. Untuk pengujian parameter secara bersama dapat digunakan uji keberartian model yaitu uji G. Statistik uji G dinotasikan sebagai berikut:

$$G = -2 \ln \left[\frac{L_0}{L_k} \right].$$

Sedangkan pengujian parameter β secara parsial dilakukan dengan membandingkan model terbaik yang dihasilkan oleh uji simultan terhadap model tanpa variabel bebas di dalam model terbaik. Statistik ujinya adalah:

$$W = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}.$$

Statistik uji Wald mengikuti sebaran normal baku Z , H_0 akan ditolak jika $W > Z_{\alpha/2}$ atau $p_{value} < \alpha$.

Estimasi model logit regresi logistik ordinal dibentuk berdasarkan pengujian secara parsial. Berdasarkan persamaan model logit, akan dilakukan prosedur pengklasifikasian status gizi balita.

Prosedur klasifikasi adalah suatu evaluasi untuk melihat peluang kesalahan klasifikasi (*misclassification*) dari fungsi klasifikasi. Prosedur klasifikasi yang baik ditentukan dengan nilai misklasifikasi yang kecil. Terdapat alat ukur yang dapat digunakan untuk menentukan kesalahan klasifikasi yaitu *apparent error rate* (APER). Penentuan kesalahan pengklasifikasian dapat diketahui melalui tabel klasifikasi (Johnson & Wichern, 2007).

Tabel 1. Tabel Klasifikasi

Keanggotaan sebenarnya	Keanggotaan prediksi			Total
	$\hat{\pi}_1$	$\hat{\pi}_2$	$\hat{\pi}_3$	
π_1	n_{11}	n_{12}	n_{13}	A
π_2	n_{21}	n_{22}	n_{23}	B
π_3	n_{31}	n_{32}	n_{33}	C
Total	D	E	F	G

Sehingga kesalahan klasifikasi dapat dirumuskan dalam:

$$APER = \left(\frac{n_{12} + n_{13} + n_{21} + n_{23} + n_{31} + n_{32}}{G} \right).$$

Nilai ketepatan klasifikasi (*correct classification*) dirumuskan dalam: $1 - APER$.

Bootstrap Aggregating (Bagging) merupakan sebuah metode yang diusulkan oleh Breiman. Metode ini digunakan sebagai alat untuk memperbaiki stabilitas dan kekuatan prediksi dengan cara mereduksi variansi dari suatu prediktor pada metode klasifikasi dan regresi yang penggunaannya tidak dibatasi hanya untuk memperbaiki estimator. Versi berganda dibentuk dengan replikasi *bootstrap* dari sebuah *data set*.

Ambil *bootstrap* sampel dengan pengulangan $\{\mathcal{L}^{(B)}\}$ dari *data set* \mathcal{L} dan membentuk $\{\varphi(x, \mathcal{L}^{(B)})\}$. Jika y merupakan data numerik, diberikan φ_B sebagai

$$\varphi_B(x) = \text{average}_B \varphi(x, \mathcal{L}^{(B)}), \quad (4)$$

Jika y merupakan kelas label, maka untuk menentukan kategori dengan melakukan *voting* atau *average* pada $\{\varphi(x, \mathcal{L}^{(B)})\}$ dari $\varphi_B(x)$. Metode *bootstrap* yang seperti ini dinamakan *bootstrap aggregating (bagging)*.

Bagging bekerja dengan baik pada metode klasifikasi yang menghasilkan parameter yang tidak stabil, yang mana perubahan kecil pada *data set* akan menghasilkan perubahan besar pada model yang diperoleh. Beberapa metode yang tidak stabil adalah *neural networks*, regresi, klasifikasi, dan regresi pohon (*regression tree*). Sedangkan contoh metode yang stabil adalah *k-nearest neighbor* (Breiman, 1996).

2. METODE PENELITIAN

Data yang digunakan pada penelitian ini adalah data sekunder dari penelitian Suciari (2015). Penelitian ini mengenai status gizi balita berdasarkan indikator tinggi badan menurut umur (TB/U) yang menggambarkan keadaan kurang gizi pendek (*stunting*), dengan objek balita umur 24-59 bulan di UPT. Puskesmas Klungkung I pada tahun 2015. Variabel respons yang digunakan pada penelitian ini terdiri atas normal, pendek, dan sangat pendek. Sedangkan variabel prediktor yang digunakan meliputi umur balita, jenis kelamin balita, panjang badan balita saat lahir, berat badan lahir, umur pemberian MP-ASI (UMPASI), status anemia

ibu saat hamil, panjang lingkaran lengan atas (LILA) ibu saat hamil, pendidikan ibu, status pekerjaan ibu, dan umur ibu.

Pada penelitian ini digunakan metode *bagging* regresi logistik ordinal dengan replikasi *bootstrap* (B). Replikasi B yang digunakan adalah 51, 61, 71, 81, 91, 101, 151, 201, 301, 401, dan 501. Adapun tahapan analisis data penelitian, yaitu:

1. Melakukan persiapan data dan membagi data menjadi data *training* dan *testing* masing-masing sebanyak 70% dan 30%.
2. Menentukan estimasi model regresi logistik ordinal pada data *training*.
3. Melakukan pengujian signifikansi parameter terhadap variabel respons secara simultan dan parsial.
4. Menentukan prediksi variabel respons dari model *bagging* regresi logistik ordinal dan menghitung tingkat ketepatan klasifikasi pada data *testing*.
5. Mengambil $\frac{1}{2}$ data (\mathcal{L}_B) dari data *training* \mathcal{L} untuk setiap replikasi sebanyak B .
6. Menentukan estimasi model *bagging* regresi logistik ordinal hasil sampel *bootstrap* \mathcal{L}_B .
7. Menentukan prediksi variabel respons dari model *bagging* regresi logistik ordinal untuk setiap replikasi menggunakan data *testing*.
8. Mengulang langkah 5 sampai langkah 7 sebanyak B kali.
9. Menentukan prediksi kelas maksimum dengan *majority voting* dan menghitung nilai kesalahan klasifikasi *bagging* (\bar{e}_B) pada data *testing*.
10. Memperoleh model *bagging* regresi logistik ordinal pada setiap replikasi B kali.
11. Membandingkan tingkat ketepatan klasifikasi model regresi logistik ordinal dan *bagging* regresi logistik ordinal.

3. HASIL DAN PEMBAHASAN

Analisis Regresi Logistik Ordinal

Pembentukan estimasi model regresi logistik ordinal, menggunakan data *training* sebanyak 70% dari 116 data status gizi balita.

Pengujian koefisien regresi secara simultan bertujuan untuk mengetahui pengaruh variabel prediktor terhadap variabel respons secara simultan (bersama-sama). Untuk pengujian parameter secara simultan dapat digunakan uji keberartian model yaitu uji G dengan hipotesis sebagai berikut:

$H_0: \beta_1 = \beta_2 = \dots = \beta_{10} = 0$ (tidak ada pengaruh sekumpulan variabel prediktor terhadap variabel respons),

$H_1: \text{minimal ada satu } \beta_i \neq 0$.

Tabel 2. Statistik Uji G

Model	-2 Likelihood Ratio	Chi-square	df	Pr (Chi)
Model1	133,5428			
Model2	102,0224	31,5204	10	0,000481

Berdasarkan nilai Pr(Chi) pada Tabel 2 sebesar $0,000481 < \alpha = 0,05$, maka keputusan tolak H_0 dan terima H_1 . Hal ini berarti terdapat minimal satu variabel respons yang berpengaruh terhadap status gizi balita.

Selanjutnya, dilakukan uji parsial. Hipotesis yang digunakan untuk pengujian secara parsial yaitu:

$H_0: \beta_i = 0$ (tidak ada pengaruh variabel bebas yang diuji terhadap variabel respons),

$H_1: \beta_i \neq 0$ (terdapat pengaruh variabel bebas yang diuji terhadap variabel respons).

Hasil pada Tabel 3. menunjukkan bahwa hasil pengujian parameter secara parsial menjelaskan variabel panjang badan lahir (X_3), berat badan lahir (X_4), dan panjang LILA ibu saat hamil (X_7) memiliki nilai signifikansi $\alpha < 0,05$ atau dengan kata lain tolak H_0 , sehingga dapat diketahui bahwa variabel-variabel tersebut memiliki pengaruh signifikan terhadap status gizi balita.

Model logit regresi logistik ordinal atau model regresi logistik ordinal berganda yang dapat dibentuk berdasarkan pengujian secara parsial adalah sebagai berikut:

$$c_1(\underline{X}) = -24,844 - 0,3181X_3 - 0,0022X_4 + 0,3442X_7, \quad (5)$$

$$c_2(\underline{X}) = -22,231 - 0,3181X_3 - 0,0022X_4 + 0,3442X_7. \quad (6)$$

Tabel 3. Pengujian Parameter Secara Parsial

Variabel	Koefisien	Std.error	t-value	p-value
X_1 : umur balita	-0,064	0,0319	-0,1991	0,8422
X_2 : jenis kelamin	-0,834	0,5608	-1,4871	0,1370
X_3 : panjang badan lahir	-0,3181	0,0909	-3,4994	0,00047
X_4 : berat badan lahir	-0,0022	0,0007	-3,3694	0,00075
X_5 : kondisi awal pemberian MP-ASI	0,9905	0,6242	1,5868	0,1126
X_6 : status anemia saat hamil	-0,2651	0,2817	-0,9410	0,3467
X_7 : panjang ingkar lengan atas (LILA) ibu saat hamil	0,3442	0,1746	1,9710	0,0487
X_8 : pendidikan ibu	-0,3969	0,3219	-1,2330	0,2176
X_9 : status pekerjaan ibu	-0,6056	0,5692	-1,0640	0,2873
X_{10} : umur ibu	0,0384	0,0470	0,8162	0,4144
Konstanta 1	-24,844	0,0013	-1874,87	0,0000
Konstanta 2	-22,231	0,5190	-42,8305	0,0000

Berdasarkan model logit regresi logistik (5) dan (6), didapatkan persamaan model regresi logistik ordinal sebagai berikut:

$$\phi_1(x) = \frac{\exp(-24,844 - 0,3181X_3 - 0,0022X_4 + 0,3442X_7)}{1 + \exp(-24,844 - 0,3181X_3 - 0,0022X_4 + 0,3442X_7)}$$

$$\phi_2(x) = \frac{\exp(-22,231 - 0,3181X_3 - 0,0022X_4 + 0,3442X_7)}{1 + \exp(-22,231 - 0,3181X_3 - 0,0022X_4 + 0,3442X_7)}$$

Pada proses pengklasifikasian digunakan data *testing* sebanyak 30% dari 116 data status gizi balita. Berikut ini merupakan hasil pengklasifikasian status gizi balita untuk kelompok normal, pendek, atau sangat pendek.

Tabel 4. Ketepatan Klasifikasi *Data set* Tunggal Status Gizi Balita

Actual Group	Prediction Group			Total
	Normal	Pendek	Sangat pendek	
Normal	21	4	1	26
Pendek	6	1	0	7
Sangat pendek	2	0	0	2
Total	29	5	1	35

Berdasarkan Tabel 4. dapat dihitung besarnya tingkat kesalahan klasifikasi:

$$APER = \left(\frac{4 + 1 + 6 + 0 + 2 + 0}{35} \right) = 0,3714$$

Dengan ketepatan klasifikasi sebesar $1 - APER = 1 - 0,3714 = 0,6286$.

Sehingga didapatkan persentase ketepatan klasifikasi berdasarkan data *testing* secara keseluruhan sebesar 62,86%. Sedangkan kesalahan klasifikasinya adalah sebesar 37,14%.

Pengklasifikasian Status Gizi Balita dengan Bagging Regresi Logistik Ordinal

Fungsi klasifikasi (5) dan (6) merupakan model *data set* tunggal. Parameter yang dihasilkan oleh model regresi logistik ordinal cenderung tidak stabil (Breiman, 1996). Untuk melihat kestabilan pada data dan memperoleh keakuratan model yang lebih baik maka digunakan *resampling* pada data. *Resampling* yang melibatkan beberapa variabel respons dan prediktor dengan metode *bagging* yang merupakan pengambilan sampel dengan pengembalian. Variabel-variabel tersebut adalah variabel pada model logit regresi logistik ordinal (5) dan (6). Sampel *bootstrap* diambil sebanyak $\frac{1}{2}$ dari data *training*. Kemudian, direplikasi *bootstrap* sebanyak 51, 61, 71, 81, 91, 101, 151, 201, 301, 401, dan 501 kali. Pada setiap replikasi *B* akan dibentuk model *bagging* regresi logistik ordinal dari setiap replikasi *bootstrap*. Penentuan prediksi untuk setiap replikasi *B* menggunakan data *testing* berdasarkan model *bagging* regresi logistik ordinal.

Majority voting digunakan untuk mendapatkan prediksi variabel respons maksimum pada replikasi B , kemudian kesalahan klasifikasi *bagging* regresi logistik ordinal ditentukan berdasarkan prediksi maksimum pada replikasi B . Sehingga, kesalahan klasifikasi *bagging* regresi logistik ordinal untuk replikasi B kali adalah \bar{e}_B .

Ketepatan Klasifikasi *Bagging* Regresi Logistik Ordinal terbesar yaitu 71,43% pada replikasi *bootstrap* 81 kali dengan misklasifikasi sebesar 28,57%. Keberhasilan *bagging* diukur dari seberapa besar *bagging* dapat menurunkan misklasifikasi dari model *data set* tunggal.

Perbandingan Tingkat Ketepatan Klasifikasi Metode Regresi Logistik Ordinal dan *Bagging* Regresi Logistik Ordinal

Perbandingan tingkat ketepatan klasifikasi metode RLO dan *Bagging* RLO dijabarkan pada Tabel 5.

Tabel 5. Perbandingan Ketepatan Klasifikasi

Ketepatan Klasifikasi RLO	<i>Bagging</i> RLO		Kenaikan Tingkat Ketepatan Klasifikasi
	Replikasi Bootstrap	Ketepatan Klasifikasi	
62,86%	51	68,57%	5,71%
	61	68,57%	5,71%
	71	68,57%	5,71%
	81	71,43%	8,57%
	91	65,71%	2,85%
	101	68,57%	5,71%
	151	68,57%	5,71%
	201	68,57%	5,71%
	301	68,57%	5,71%
	401	68,57%	5,71%
501	68,57%	5,71%	

Berdasarkan Tabel 5 diketahui bahwa persentase tingkat ketepatan klasifikasi *bagging* regresi logistik ordinal terbesar pada replikasi *bootstrap* 81 kali yaitu sebesar 71,43%. Persentase tingkat ketepatan klasifikasi menunjukkan kestabilan pada replikasi *bootstrap* 51, 61, 71, 101, 151, 201, 301, 401, dan 501 kali, yaitu sebesar 68,57% dengan misklasifikasi sebesar 31,43%. Dengan demikian, dapat diketahui bahwa ukuran *bootstrap* B tidak berpengaruh terhadap tingkat ketepatan klasifikasi, namun menunjukkan

kekonvergenan tingkat ketepatan klasifikasi. Keberhasilan *bagging* diukur dari seberapa besar *bagging* dapat menurunkan misklasifikasi dari model *data set* tunggal. Penurunan misklasifikasi dapat diukur dengan membandingkan tingkat ketepatan klasifikasi *bagging* regresi logistik ordinal dan regresi logistik ordinal.

Berikut ini merupakan nilai AIC *bagging* regresi logistik ordinal yang dibentuk berdasarkan pada replikasi *bootstrap* pada replikasi B kali, yaitu:

Tabel 6. Nilai AIC *Bagging* Regresi Logistik Ordinal

Replikasi	AIC	Replikasi	AIC
51	57,722	151	61,657
61	61,867	201	60,309
71	58,219	301	50,976
81	56,582	401	48,886
91	52,049	501	48,551
101	60,159		

Model terbaik dapat ditentukan dengan melihat nilai AIC terkecil. Berdasarkan model logit *bagging* regresi logistik ordinal pada Tabel 6 dapat diketahui pada replikasi 501 kali memiliki nilai AIC yang lebih kecil dibandingkan dengan replikasi B lainnya. Hal ini berarti bahwa model tersebut merupakan model terbaik. Dengan demikian, replikasi 501 kali lebih tepat digunakan untuk mengklasifikasikan status gizi balita di Kabupaten Klungkung.

Interpretasi Koefisien *Bagging* Regresi Logistik Ordinal

Estimasi model logit *bagging* regresi logistik ordinal pada replikasi *bootstrap* 501 kali yang diperoleh adalah sebagai berikut:

$$c_1(\underline{X}) = -32,797 - 0,500 X_3 - 0,0035 X_4 + 0,419X_7,$$

$$c_2(\underline{X}) = -30,848 - 0,500 X_3 - 0,0035 X_4 + 0,419X_7.$$

Sehingga estimasi model *bagging* regresi logistik yang dapat dibentuk adalah:

$$\phi_1(x) = \frac{\exp(-32,797 - 0,500 X_3 - 0,0035 X_4 + 0,419X_7)}{1 + \exp(-32,797 - 0,500 X_3 - 0,0035 X_4 + 0,419X_7)}$$

$$\phi_2(x) = \frac{\exp(-30,848 - 0,500 X_3 - 0,0035 X_4 + 0,419 X_7)}{1 + \exp(-30,848 - 0,500 X_3 - 0,0035 X_4 + 0,419 X_7)}$$

Pada model estimasi logit *bagging* regresi logistik ordinal, yang digunakan untuk interpretasi koefisien adalah *odds ratio*. Nilai *odds ratio* merupakan rasio antara kecenderungan (risiko) terjadinya suatu peristiwa. Nilai *odds ratio* pada *bagging* regresi logistik ordinal dapat dilihat dari nilai $\exp(\beta_i)$.

Nilai *odds ratio* masing-masing koefisien pada model terbaik yang diperoleh yaitu model status gizi balita yang dipengaruhi oleh variabel panjang badan lahir (X_3), berat badan lahir (X_4), dan panjang LILA ibu saat hamil (X_7) seperti tampak pada Tabel 7.

Tabel 7. Nilai *Odds Ratio* masing-masing Variabel

Variabel	Koefisien	$\exp(\beta_i)$
X_3 : panjang badan lahir	-0,500	0,607
X_4 : berat badan lahir	-0,0035	0,997
X_7 : panjang lingkaran lengan atas (LILA) ibu saat hamil	0,419	1,520

Berdasarkan Tabel 7 dapat diinterpretasikan sebagai berikut:

- i. *Odds ratio* X_3 (panjang badan lahir) = 0,607, artinya dengan kenaikan panjang badan lahir sebesar 1 cm, akan menurunkan risiko tingkat status gizi balita sangat pendek sebesar 0,607 kali.
- ii. *Odds ratio* X_4 (berat badan lahir) = 0,997, artinya setiap peningkatan berat badan lahir sebesar 1 gram, maka risiko menurunnya tingkat status gizi balita sangat pendek adalah sebesar 0,997 kali.
- iii. *Odds ratio* X_7 (panjang LILA ibu saat hamil) = 1,520, artinya ibu hamil yang memiliki LILA dengan kategori normal berisiko memiliki balita berstatus gizi normal lebih besar dibandingkan dengan ibu hamil yang memiliki LILA dengan KEK atau dengan kata lain, risiko balita memiliki status gizi sangat pendek akan menurun sebesar 1,520 kali.

4. KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil dan pembahasan diperoleh model logit *bagging* regresi logistik ordinal, dan tiga variabel prediktor yang signifikan berpengaruh terhadap variabel respons, yaitu tinggi badan lahir (X_3), badan lahir (X_4), dan panjang LILA ibu saat hamil (X_7).

Tingkat ketepatan klasifikasi pada data *testing* dengan metode regresi logistik ordinal sebesar 62,86%. Tingkat ketepatan klasifikasi dengan metode *bagging* regresi logistik ordinal untuk mengklasifikasikan status gizi balita umur 24-59 bulan di UPT. Puskesmas Klungkung I pada replikasi 501 kali dapat meningkatkan ketepatan klasifikasi dari model regresi logistik ordinal atau dengan kata lain *bagging* dapat memperlihatkan peningkatan ketepatan klasifikasi sebesar 5,71%. Hal ini berarti, *bagging* regresi logistik ordinal dapat mengklasifikasikan status gizi balita lebih baik dibandingkan dengan regresi logistik ordinal.

Saran

Pada penelitian dengan masalah pengklasifikasian dapat digunakan metode *bagging* untuk menaikkan tingkat ketepatan klasifikasi. Untuk ibu hamil dianjurkan menjaga panjang lingkaran lengan atas (LILA) pada masa kehamilan tetap normal, yaitu pada ukuran $LILA \geq 23,5$ cm.

DAFTAR PUSTAKA

- Akbar, M. S., Mukarromah, A., dan Paramita. 2010. Klasifikasi Status Gizi Balita dengan *Bagging* Regresi Logistik Ordinal. *Media Statistika*, Volume 3, No. 2, Hal. 103-114.
- Breiman, L., 1996. *Bagging Predictors*. Technical University of California.
- Dinas Kesehatan Provinsi Bali. 2014. Profil Kesehatan Provinsi Bali Tahun 2013. Bali: Dinas Kesehatan Provinsi Bali.
- Hosmer, D. W., and Lemeshow, S., 2000. *Applied Logistic Regression*. New York: John Wiley & Sons, Inc.

- Insani, Fitrah., AF, Saleh., dan LP, Talangko. 2015. Metode *Bootstrap Aggregating* Regresi Logistik untuk Peningkatan Ketepatan Klasifikasi Regresi Logistik Ordinal. *Jurnal Statistika FMIPA UNHAS*.
- Johnson, R. A., and Winchern, D.W., 2007. *Applied Multivariate Statistical Analysis*. USA: Prentice Hall. Inc
- Suciari, L.S. 2015. Hubungan antara Status Gizi Ibu saat Hamil, Panjang Badan, Berat Badan Lahir dan Umur Awal Pemberian MP-ASI. Denpasar: *Tugas Akhir* Universitas Udayana.
- Kementerian Kesehatan. 2013. Riset Kesehatan Dasar, RISKESDAS dalam Angka Provinsi Bali 2013. Jakarta: Lembaga Penerbitan Badan Litbangkes Kementerian Kesehatan RI.
- Subianto, Muhammad. 2009. *Understanding Classification*. Universitas Utrecht.