
PEMODELAN MARS DAN REGRESI LOGISTIK RUMAH TANGGA MISKIN KALIMANTAN TENGAH TAHUN 2016
(MARS AND LOGISTIC REGRESSION MODELING FOR HOUSEHOLD POVERTY IN CENTRAL BORNEO 2016)

Ananto Wibowo¹, Ellen Lelian Mehrani²

¹BPS Kotawaringin Barat, ananto.wibowo@bps.go.id

²BPS Kalimantan Tengah, ellen.lijian@bps.go.id

Abstrak

Kemiskinan merupakan masalah multidimensi yang dapat menyebabkan berbagai persoalan baik secara ekonomi, sosial budaya dan politik. Oleh karena itu, pemerintah berupaya merumuskan kebijakan yang sesuai untuk mengatasi hal tersebut. Upaya pemerintah perlu didukung dengan analisis ekonomi menggunakan model statistik yang akurat. Studi ini bertujuan untuk menganalisis variabel-variabel yang signifikan berpengaruh terhadap rumah tangga miskin Kalimantan Tengah dengan regresi logistik dan *Multivariate Adaptive Regression Splines* (MARS) serta membandingkan performa kedua metode tersebut. Data yang digunakan adalah data SUSENAS Tahun 2016 bersumber dari Badan Pusat Statistik (BPS). Hasil kedua metode menunjukkan bahwa variabel jumlah anggota rumah tangga memiliki peran terbesar sebagai penentu kemiskinan rumah tangga. Selain itu, regresi logistik juga dianggap memiliki performa yang lebih baik dalam pemodelan dibandingkan metode MARS karena mempunyai nilai APER yang lebih kecil dan jumlah variabel signifikan yang lebih banyak.

Kata kunci: Kemiskinan, Regresi Logistik, MARS

Abstract

Poverty is a multidimensional issues that can cause various problems in economic, social, culture and politics. Therefore, the government tries to make suitable policy to solve the problem. To make a good decision, an economic analysis with accurate statistics model is needed to be done. This study aims to analyze variables that influence poor household in Kalimantan Tengah by using two methods, Logistic Regression and Multivariate Adaptive Regression Splines (MARS) besides try to find the good model by comparing APER and the amount of significant variables from those two methods. The source of data was 2016 Socioeconomic Survey (SUSENAS) from Central Bureau of Statistics. Based on those methods, the amount of people in a household has the biggest role in household poverty. This research also shows that Logistic Regression method has better perform than MARS method because logistic regression has smaller APER value and bigger amount of significant variables.

Keywords: Poverty, Logistic Regression, MARS

PENDAHULUAN

Masalah kemiskinan merupakan salah satu persoalan mendasar yang menjadi pusat perhatian pemerintah di seluruh wilayah termasuk Provinsi Kalimantan Tengah (Kalteng). Upaya pemerintah Kalteng dalam pengentasan kemiskinan masih perlu ditingkatkan. Meski persentase penduduk miskin terus menurun, tetap perlu perhatian khusus untuk pemerataan ekonomi. Data statistik telah menunjukkan bahwa persentase penduduk miskin Kalteng masih sebesar 5,66 persen (Badan Pusat Statistik, 2016). Oleh karena itu, perlu adanya suatu kajian yang berkenaan dengan fenomena tingkat kemiskinan di Kalteng demi keberhasilan pembangunan. Kajian yang dimaksud dapat menggunakan model analisis secara mikro rumah tangga karena variabelnya memiliki peran penting dalam menganalisis determinan rumah tangga miskin. Variabel tersebut terdiri dari variabel demografi, sosial, dan ekonomi.

Berdasarkan data dari Badan Pusat Statistik (2008), beberapa variabel mikro yang memiliki keterkaitan erat dengan kemiskinan secara konseptual yaitu jumlah anggota rumah tangga, kepala rumah tangga perempuan, sumber penghasilan utama rumah tangga, status pekerjaan, luas lantai rumah perkapita, jenis lantai rumah, jenis atap rumah, jenis dinding rumah, sumber penerangan rumah, akses terhadap air bersih, fasilitas jamban dan status kepemilikan rumah. Pada skala mikro, kajian lebih menekankan pada status rumah tangga miskin dan tidak miskin yang merupakan respon biner (kategorik).

Dalam ilmu statistika, terdapat beberapa metode yang dapat digunakan untuk mengetahui pengaruh variabel prediktor terhadap variabel respon berupa kategorik seperti metode regresi logistik dan *Multivariate Adaptive Regressions Spline* (MARS). Kelebihan regresi logistik adalah tidak menghendaki adanya pemenuhan asumsi (Hair, Black, Babin, & Anderson, 2010). Selain itu regresi logistik digunakan untuk data non linear dan mudah diinterpretasikan. Namun, multikolinieritas antarvariabel prediktor menjadi suatu masalah dan membuat estimasi parameter menjadi tidak stabil (Thompson, 2009). Metode MARS mampu mengakomodir interaksi antarvariabel bahkan yang berdimensi tinggi (Mina, & Barrios, 2010). Meski demikian metode MARS memiliki kelemahan diantaranya bentukan model yang sangat kompleks karena interaksi yang terjadi. Selain itu, metode MARS seringkali tidak dapat menangkap efek koefisien masing-masing variabel prediktor terhadap variabel respon (Melati, 2011).

Pada penelitian ini terdapat dua tujuan yang hendak dicapai, diantaranya menganalisis variabel-variabel mikro yang berpengaruh signifikan terhadap status kemiskinan rumah tangga Provinsi Kalimantan Tengah Tahun 2016 dan mengetahui regresi mana yang memiliki performa terbaik dalam pengklasifikasian dengan metode MARS dan regresi logistik.

METODE

Regresi Logistik

Regresi logistik menjelaskan model matematis antara variabel prediktor terhadap variabel respon yang bersifat biner. Variabel respon dikategorikan sebagai $Y = 1$ pada kategori sukses dan $Y = 0$ untuk kategori gagal (Hosmer & Lemeshow, 2000). Nilai rata-rata dari variabel respon regresi logistik diperoleh dari nilai variabel penjelas dengan *conditional mean* yang dinotasikan sebagai $E(Y|x)$ dengan Y merupakan variabel respon dan x merupakan variabel penjelas.

Nilai *conditional mean* harus bernilai lebih besar atau sama dengan nol dan kurang dari atau sama dengan 1. $E(Y|x)$ dinotasikan sebagai $\pi(x)$ untuk menyatakan nilai *conditional mean* dari Y yang didapatkan dari x yang ditetapkan ketika distribusi logistik digunakan (Hosmer & Lemeshow, 2000) atau peluang terjadinya kejadian sukses yaitu $y = 1$ pada suatu nilai x (Agresti, 2007), sehingga model regresi logistik dapat diformulasikan sebagai berikut.

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}} \quad (1)$$

dengan β_p adalah nilai parameter ke- j , dimana $j = 0, 1, \dots, p$, dan p adalah jumlah variabel.

Pengujian parameter secara simultan dilakukan dengan menggunakan *Likelihood Ratio Test* (Hosmer & Lemeshow, 2000). Hipotesis yang digunakan yaitu:

H_0 : $\beta_1 = \beta_2 = \dots = \beta_p = 0$ (tidak terdapat pengaruh variabel penjelas secara simultan terhadap variabel respon)

H_1 : minimal terdapat satu $\beta_j \neq 0$; $j = 1, 2, \dots, p$ (minimal terdapat satu variabel penjelas yang berpengaruh terhadap variabel respon)

Statistik uji yang digunakan yaitu statistik uji G dengan rumusan sebagai berikut.

$$G = -2 \ln \left[\frac{L_0}{L_1} \right] \sim \chi^2_{(p)} \quad (2)$$

Jika hasil pengujian simultan menghasilkan keputusan H_0 ditolak, maka dilanjutkan pengujian parameter secara parsial. Pengujian parameter secara parsial atau uji parsial digunakan untuk mengetahui pengaruh tiap-tiap variabel penjelas terhadap variabel respon. Pengujian dapat dilakukan dengan menggunakan uji *Wald* (Hosmer & Lemeshow, 2000). Hipotesis yang digunakan yaitu:

H_0 : $\beta_j = 0$ (tidak terdapat pengaruh variabel penjelas ke- j terhadap variabel respon)

H_1 : $\beta_j \neq 0$ (terdapat pengaruh variabel penjelas ke- j terhadap variabel respon)

Statistik uji yang digunakan adalah:

$$W_j = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right)^2 \sim \chi^2_{(1)} \quad (3)$$

dimana,

j : 1, 2, ..., p

W_j : statistik uji Wald untuk variabel penjelas ke- j

$\hat{\beta}_j$: penduga parameter β_j

$SE(\hat{\beta}_j)$: standar error dari $\hat{\beta}_j$

H_0 akan ditolak apabila $W_j > \chi^2_{(1, \alpha)}$ atau $p\text{-value} < \alpha$ (Hosmer & Lemeshow, 2000). Jika H_0 ditolak dapat disimpulkan bahwa terdapat pengaruh variabel penjelas ke- j terhadap variabel respon.

Multivariate Adaptive Regression Splines (MARS)

Model MARS merupakan kombinasi yang kompleks dari *Recursive Partitioning Regression* dan *Spline*. Regresi *spline* menghasilkan persamaan bentuk parametrik polinomial *piecewise* (terbagi dalam beberapa region), adapun RPR merupakan salah satu pendekatan komputasi yang digunakan untuk data berdimensi tinggi. Metode MARS digunakan pada data yang berdimensi tinggi

dimana variabel prediktornya dibatasi 3 hingga 20 serta memiliki ukuran sampel yang berukuran ideal dari 50 hingga 2000 (Friedman, 1990).

Dalam pemodelan MARS dilakukan *trial and error* terhadap kombinasi jumlah *Basis Function* (BF), *Maksimum Interaktion* (MI) dan *Minimum Observation* (MO) pada knot untuk mendapatkan model MARS terbaik. Jumlah BF yang dipakai 2 hingga 4 kali banyaknya variabel prediktor. Adapun jumlah interaksi yang disarankan adalah 1 hingga 3 karena interpretasi akan sulit dilakukan jika lebih dari 3 interaksi. Sedangkan jumlah minimum observasi setiap knot yaitu 0, 1, 2 dan 3 dimana semakin besar MO maka akan semakin mengurangi fleksibilitas model dan akurasi (Febrianto, 2010).

Estimator model MARS dapat ditulis sebagai berikut.

$$\hat{f}(x) = a_0 + \sum_{m=1}^M a_m \prod_{k=1}^{K_m} [s_{km}(x_{v(k,m)} - t_{km})] \quad (4)$$

Keterangan:

- a_0 : konstanta
- a_m : koefisien basis fungsi ke-m
- M : maksimum basis fungsi
- K_m : derajat interaksi
- s_{km} : nilainya ± 1
- $x_{v(k,m)}$: variabel prediktor/independen
- t_{km} : nilai knots dari variabel prediktor/independent
- q : derajat *spline*

Pada model MARS, selain penentuan knots yang didasarkan pada data dan menghasilkan fungsi yang kontinu, pemilihan model terbaik pada model MARS juga menggunakan algoritma *stepwise (forward dan backward)*. Model yang terbaik memiliki nilai GCV yang minimum dibandingkan dengan model-model lain. Nilai GCV didefinisikan dengan formula:

$$GCV(M) = \frac{ASR}{\left[1 - \frac{\tilde{C}(M)}{N}\right]^2} = \frac{\frac{1}{N} \sum_{i=1}^N [y_i - \hat{f}_M(x_i)]^2}{\left[1 - \frac{\tilde{C}(M)}{N}\right]^2} \quad (5)$$

Keterangan:

- x_i : variabel prediktor
- y_i : variabel respon
- $\tilde{C}(M)$: $C(M) + dM$
- $C(M)$: trace $[\mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T] + 1$
- N : banyaknya pengamatan

Tabel dan Performa Klasifikasi

Tabel klasifikasi merupakan hasil klasifikasi silang terhadap variabel respon Y yang berbentuk dua kategori (Hosmer & Lemeshow, 2000). Tabel klasifikasi ditentukan oleh nilai *cut point* sebagai pembanding dari masing-masing estimasi peluang regresi logistik. Jika estimasi peluang observasi lebih dari *cut point* maka observasi dimasukkan terhadap kategori 1 (sukses), sedangkan untuk yang kurang dari *cut point* dimasukkan ke dalam kategori 0 (gagal). Pada penelitian ini nilai *cut point* digunakan adalah 0,5. Baik model MARS maupun Regresi Logistik, keduanya merupakan model klasifikasi sehingga ketentuan *cut point* yang sama dapat dilakukan pada kedua model regresi tersebut.

Terdapat beberapa ukuran penting yang dapat mengukur performa pengklasifikasian suatu model. Pada penelitian ini digunakan *Apparent Error Rate*

(APER) yang merupakan tingkat kesalahan klasifikasi dimana nilai yang semakin kecil maka model terbentuk akan semakin baik.

Tabel 1. Tabel Klasifikasi

Hasil Observasi (Actual Class)	Taksiran (Predicted Class)	
	Rumah Tangga tidak miskin (0)	Rumah Tangga Miskin (1)
Rumah Tangga Tidak Miskin (0)	n_{11}	n_{12}
Rumah Tangga Miskin (1)	n_{21}	n_{22}

Berdasarkan Tabel 1, formula APER (Timm, 2002) dapat dirumuskan sebagai berikut:

$$APER = \frac{n_{12} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}} \times 100\% \quad (6)$$

Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2016. Pada data tersebut sudah terklasifikasi rumah tangga miskin dan tidak miskin berdasarkan garis kemiskinan. Demi keperluan analisis, terdapat enam variabel yang digunakan dalam penelitian diantaranya adalah status kemiskinan rumah tangga, jumlah anggota rumah tangga (ART), jenis kelamin kepala rumah tangga (KRT), pendidikan KRT, sektor pekerjaan KRT, dan status pekerjaan KRT. Keenam variabel tersebut dirinci pada Tabel 2.

Tabel 2. Variabel Penelitian

No	Variabel	Kategori
1.	Status Kemiskinan Rumah Tangga (Y)	0 = Tidak miskin 1 = Miskin
2.	Jumlah ART (X_1)	0 = Jumlah ART \leq 4 1 = Jumlah ART $>$ 4
3.	Jenis Kelamin KRT (X_2)	0 = Laki-Laki 1 = Perempuan
4.	Pendidikan KRT (X_3)	0 = Minimal Tamat SMA 1 = Maksimal Tamat SMP
5.	Sektor Pekerjaan KRT (X_4)	0 = Non Pertanian 1 = Pertanian
6.	Status Pekerja KRT (X_5)	0 = Formal 1 = Informal

HASIL DAN PEMBAHASAN

Model Regresi Logistik

Model regresi logistik rumah tangga miskin yang terbentuk berdasarkan hasil pengolahan adalah sebagai berikut:

$$\pi(x) = \frac{e^{-4,797 + 1,583X_1 + 0,856X_2 + 0,511X_3 + 0,344X_4 + 0,535X_5}}{1 + e^{-4,797 + 1,583X_1 + 0,856X_2 + 0,511X_3 + 0,344X_4 + 0,535X_5}}$$

Apabila persamaan dilakukan transformasi logit maka regresi logistik akan memiliki persamaan :

$$\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = -4,797 + 1,583X_1 + 0,856X_2 + 0,511X_3 + 0,344X_4 + 0,535X_5$$

Hasil uji simultan *Likelihood Ratio Test* menyimpulkan bahwa minimal terdapat salah satu variabel prediktor yang berpengaruh terhadap variabel respon. Hal ini dapat ditunjukkan oleh nilai *G* sebesar 1873,166 dengan taraf signifikansi (sig) 0,000 yang berarti H_0 ditolak.

Selanjutnya, dilakukan uji secara parsial dengan menggunakan Uji Wald. Berdasarkan hasil uji parsial, seluruh variabel yang diolah telah signifikan dalam memengaruhi status rumah tangga miskin. Hal ini ditunjukkan oleh nilai kolom signifikansi Tabel 3, dimana masing-masing variabel memiliki nilai yang mendekati nol atau berkesimpulan bahwa H_0 ditolak.

Tabel 3. Hasil Uji Parsial Model Regresi Logistik

Variabel	B	S.E.	Wald	df	Sig	Exp(B)
X_1	1,583	0,136	135,330	1	,000	4,870
X_2	0,856	0,198	18,598	1	,000	2,354
X_3	0,511	0,176	8,443	1	,004	1,667
X_4	0,344	0,143	5,785	1	,016	1,411
X_5	0,535	0,156	11,732	1	,001	1,707

Adapun nilai *Odds Ratio* juga ditunjukkan oleh kolom Exp(B) yang berguna untuk interpretasi model regresi sebagai berikut.

1. Variabel X_1 : suatu rumah tangga yang memiliki jumlah ART lebih dari empat orang memiliki kecenderungan miskin sebesar 4,870 kali dibandingkan rumah tangga yang memiliki jumlah ART kurang dari sama dengan empat orang. Ukuran rumah tangga dalam analisis kemiskinan berkaitan dengan besarnya tanggungan yang dimiliki suatu rumah tangga. Rumah tangga miskin cenderung memiliki jumlah ART yang lebih banyak. Karena rumah tangga miskin cenderung mempunyai tingkat kelahiran yang tinggi. Tingkat kematian anak pada rumah tangga miskin juga relatif tinggi akibat kurangnya pendapatan dan akses kesehatan serta pemenuhan gizi anak. Pada Tabel 3, variabel jumlah ART ini memiliki nilai koefisien terbesar atau dengan kata lain efek variabel jumlah ART merupakan yang tertinggi terhadap penentu status rumah tangga miskin.
2. Variabel X_2 : suatu rumah tangga yang berkepala rumah tangga perempuan memiliki kecenderungan miskin sebesar 2,354 kali dibandingkan rumah tangga yang berkepala rumah tangga laki-laki. Secara umum peran perempuan sebagai kepala rumah tangga dalam memenuhi kebutuhan hidup keluarga akan mengalami banyak kendala dibandingkan laki-laki. Status KRT perempuan bisa saja belum kawin atau cerai yang memiliki dampak terhadap pendapatan karena jumlah anggota yang bekerja hanya seorang.
3. Variabel X_3 : suatu rumah tangga yang pendidikan kepala rumah tangganya maksimal tamat SMP memiliki kecenderungan miskin sebesar 1,667 kali dibandingkan rumah tangga yang pendidikan kepala rumah tangganya minimal tamat SMA. Hubungan antara kemiskinan dan pendidikan sangatlah penting karena pendidikan sangat berperan dalam memengaruhi kemiskinan. Pendidikan merupakan indikator penentu kualitas sumber daya.

Dengan pendidikan seseorang bisa berpikir, memiliki pengetahuan dan membangun kepercayaan diri.

4. Variabel X_4 : suatu rumah tangga yang sektor pekerjaan utama kepala rumah tangganya bergantung pada sektor pertanian memiliki kecenderungan miskin sebesar 1,411 kali dibandingkan rumah tangga yang sektor pekerjaan utama kepala rumah tangganya bergantung pada sektor nonpertanian. Profil orang miskin seringkali melekat dengan mereka yang bekerja di sektor pertanian seperti petani gurem, nelayan, buruh tani dan perkebunan serta pencari kayu dan madu di hutan.
5. Variabel X_5 : suatu rumah tangga yang status pekerjaan utama kepala rumah tangganya pada sektor informal memiliki kecenderungan miskin sebesar 1,707 kali dibandingkan rumah tangga yang yang status pekerjaan utama kepala rumah tangganya pada sektor formal. Secara umum, KRT yang berstatus sebagai berusaha sendiri atau berusaha dibantu buruh tidak tetap/tidak dibayar cenderung memiliki proporsi untuk menjadi rumah tangga miskin yang lebih besar dibanding status pekerja lainnya.

Selain mengetahui variabel-variabel apa saja yang berpengaruh terhadap rumah tangga miskin, regresi logistik juga dapat mengetahui klasifikasi rumah tangga miskin. Pada klasifikasi tersebut dapat diketahui seberapa besar keakuratan pengklasifikasian dengan menggunakan regresi logistik.

Tabel 4. Klasifikasi Regresi Logistik

Hasil Observasi (<i>Actual Class</i>)	Taksiran (<i>Predicted Class</i>)		Total
	Rumah Tangga tidak miskin (0)	Rumah Tangga Miskin (1)	
Rumah Tangga Tidak Miskin (0)	5982	0	5982
Rumah Tangga Miskin (1)	246	0	246
Total	6228	0	6228

Tabel 4 memberikan informasi mengenai nilai APER hasil pengklasifikasian dengan pendekatan regresi logistik. Berdasarkan persamaan (6) maka diperoleh nilai APER sebesar 3,94 persen. Tingkat kesalahan yang ditunjukkan oleh nilai APER tergolong sangat kecil karena masih berada pada rentang kurang dari lima persen.

Model MARS

Berdasarkan hasil kombinasi pada Lampiran 2 MO, MI dan BF didapat bahwa model ke 5 dengan kombinasi BF=10, MI=2, MO=0 dianggap paling baik karena memiliki GCV minimum dan *total corrected* tertinggi. Model MARS tersebut dapat dituliskan sebagai berikut:

$$\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = 0.153 - 0.072*BF1 - 0.037*BF3 - 0.040*BF5 - 0.025*BF$$

dengan:

- BF1 = $(X_1=0)$
- BF3 = $(X_5=0)$
- BF4 = $(X_5=1)$

- BF5 $= (X_2=0)$
- BF7 $= (X_4=0) * BF4$

Hasil model MARS menunjukkan bahwa model terbaik memiliki empat basis fungsi dimana tiga basis fungsi tanpa interaksi (satu variabel) dan satu basis dengan interaksi dua variabel. Adanya interaksi antara status pekerja KRT dan sektor pekerjaan KRT terhadap rumah tangga miskin menunjukkan bahwa terdapat efek *cross product* dari beberapa buah variabel prediktor terhadap variabel respon (Nash & Bradford, 2001).

Selanjutnya, hasil GCV model MARS menyimpulkan bahwa jumlah ART memiliki tingkat kepentingan 100 persen terhadap status kemiskinan rumah tangga. Disusul oleh status pekerjaan KRT, jenis kelamin dan sektor pekerjaan KRT dengan tingkat kepentingan masing-masing 41,40 persen, 26,98 persen, dan 20,01 persen. Sedangkan pendidikan KRT tidak signifikan dalam memengaruhi status kemiskinan rumah tangga.

Berdasarkan uji simultan, model MARS yang terbentuk telah signifikan pada taraf nyata 0,05 dengan nilai *p-value* $0,999 \times 10^{-15}$. Dengan demikian keputusannya adalah tolak H_0 . Dapat disimpulkan bahwa secara umum model MARS yang terpilih dapat menggambarkan hubungan keseluruhan variabel prediktor dan variabel respon.

Apabila ditinjau pada uji parsial semua fungsi basis memiliki pengaruh yang signifikan terhadap model. Lampiran 3 pada tabel *ordinary least square* memberikan informasi mengenai uji simultan dan uji parsial dari output model MARS.

Tabel 5. Nilai Koefisien dan Odds Ratio Setiap Fungsi Basis

No	Basis Fungsi	Koefisien	Odds Ratio
1	BF1	-0,072	0,93
2	BF3	-0,037	0,96
3	BF5	-0,040	0,96
4	BF7	-0,025	0,98

Untuk mengetahui lebih jauh efek prediktor dalam memengaruhi variabel respon, diperlukan interpretasi koefisien fungsi basis dari model. Dikarenakan pemodelan MARS dalam penelitian ini merupakan respon biner, maka interpretasi dilakukan dengan menggunakan *odds ratio* yang ditunjukkan oleh Tabel 5. Interpretasi fungsi basis, yang terdiri dari dua yaitu fungsi basis satu variabel tanpa interaksi dan fungsi basis dua variabel, adalah sebagai berikut:

1. *BF1*: suatu rumah tangga yang memiliki jumlah ART kurang dari sama dengan empat orang memiliki kecenderungan miskin sebesar 0,93 kali dibandingkan rumah tangga yang memiliki jumlah ART lebih dari empat orang.
2. *BF3*: suatu rumah tangga yang memiliki KRT bekerja pada sektor formal memiliki kecenderungan miskin sebesar 0,96 kali dibandingkan rumah tangga yang memiliki KRT bekerja pada sektor informal.
3. *BF5*: suatu rumah tangga yang memiliki KRT berjenis kelamin laki-laki memiliki kecenderungan miskin sebesar 0,96 kali dibandingkan rumah tangga yang memiliki KRT berjenis kelamin perempuan.

4. *BF7*: suatu rumah tangga yang memiliki KRT bekerja pada sektor pertanian dan formal memiliki kecenderungan miskin sebesar 0,98 kali dibandingkan rumah tangga yang memiliki KRT bekerja pada sektor non pertanian dan informal.

Kemudian, analisis model MARS yang berkaitan dengan klasifikasi disajikan pada Tabel 6. Perhitungan nilai APER berdasarkan persamaan (6) sebesar 17,07 persen. Tingkat kesalahan ini tergolong cukup besar karena berada di rentang lebih dari lima persen.

Tabel 6. Klasifikasi MARS

Hasil Observasi (<i>Actual Class</i>)	Taksiran (<i>Predicted Class</i>)		Total
	Rumah Tangga tidak miskin (0)	Rumah Tangga Miskin (1)	
Rumah Tangga Tidak Miskin (0)	5054	928	5982
Rumah Tangga Miskin (1)	135	111	246
Total	6228	1039	6228

Performa Model Regresi Logistik dan MARS

Untuk keterbandingan hasil model regresi logistik dan MARS digunakan nilai APER serta jumlah variabel yang signifikan dalam penentuan status rumah tangga miskin. Idealnya suatu model memiliki nilai APER yang kecil dan jumlah variabel signifikan yang terbanyak. Berdasarkan Tabel 7, tingkat kesalahan klasifikasi model regresi logistik lebih kecil daripada model MARS yaitu 3,94 persen berbanding 17,07 persen, sedangkan bila ditinjau dari jumlah variabel yang signifikan, model regresi logistik lebih banyak dari model MARS. Dengan demikian, model status kemiskinan rumah tangga di Provinsi Kalteng dengan model regresi logistik lebih baik daripada model MARS karena memiliki nilai APER yang jauh lebih kecil dan jumlah variabel yang signifikan lebih banyak.

Tabel 7. Nilai APER dan Jumlah Variabel dari Regresi Logistik dan MARS

Metode	APER	Jumlah Variabel
Regresi Logistik	3,94 %	5
MARS	17,07 %	4

KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan maka dapat disimpulkan bahwa seluruh variabel pada regresi logistik signifikan terhadap status rumah tangga miskin Provinsi Kalteng. Variabel tersebut diantaranya jumlah ART, jenis kelamin, pendidikan KRT, sektor pekerjaan utama KRT, dan status pekerjaan utama KRT. Sedangkan pada model MARS ada satu variabel yang tidak signifikan di dalam model yaitu pendidikan KRT. Kedua model juga menyatakan bahwa jumlah ART memiliki peran terbesar sebagai penentu rumah tangga miskin di Provinsi Kalteng. Namun demikian, model regresi logistik lebih sesuai dalam

menggambarkan status rumah tangga miskin Provinsi Kalteng dibandingkan model MARS karena memiliki nilai APER yang lebih kecil dan jumlah variabel signifikan yang lebih banyak.

Saran bagi pemerintah yaitu sebaiknya program KB yang telah dicanangkan mampu lebih menjangkau ke seluruh masyarakat serta wilayah di Kalteng. Perhatian ini perlu diutamakan mengingat jumlah anggota rumah tangga memiliki andil besar dalam penentu rumah tangga miskin. Adapun saran untuk penelitian selanjutnya bisa mempertimbangkan variabel-variabel tambahan seperti proporsi ART yang bekerja, variabel perumahan, variabel usia KRT, serta variabel lainnya agar model kemiskinan yang terbentuk dapat lebih baik dengan pemodelan kategorik lainnya seperti pohon klasifikasi dan analisis diskriminan.

DAFTAR RUJUKAN

- Agresti, A. (2007). *An introduction to categorical data analysis second edition*. Florida: A John Wiley & Sons, Inc.
- Badan Pusat Statistik. (2008). *Analisis dan perhitungan tingkat kemiskinan 2008*. Jakarta. <https://doi.org/Katalog BPS 3205015>
- Badan Pusat Statistik. (2016). *Persentase penduduk miskin kalimantan tengah*. <https://kalteng.bps.go.id/statictable/2017/06/22/424/persentase-penduduk-miskin-provinsi-kalimantan-tengah-1999-2017.html>
- Friedman, H. (1990). Multivariate adaptive regression splines. *The Analysis of Statistics*, 19(1) : 1-67.
- Febrianto, A. R. (2010) *Determinan ketahanan pangan tingkat kabupaten di indonesia tahun 2007 (pendekatan multivariate adaptive regression spline)*. Jakarta: Sekolah Tinggi Ilmu Statistik.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*. <https://doi.org/10.1016/j.ijpharm.2011.02.019>
- Hosmer, D. & Lemeshow, S. (2000). *Applied logistic regression*. New York: John Wiley & Sons, Inc.
- Melati, Suci (2011). *Faktor-faktor yang mempengaruhi morbiditas malaria di provinsi papua dan papua barat tahun 2010 dengan menggunakan multivariate adaptive regression splines (mars)*. Jakarta: Sekolah Tinggi Ilmu Statistik.
- Mina, C. D., & Barrios, E. B. (2010). Profiling poverty with multivariate adaptive regression splines. *Phillipine Journal of Development*, 37(2): 55-97.
- Nash, M. & Bradford, D. (2001). *Parametric and nonparametric logistic regressions for prediction of presence / absence of an amphibian*. United States : Office of Research and Development Washington DC.
- Thompson, W. (2009). *Variable selection of correlated predictors in logistic regression: investigating the diet-heart hypothesis*. Florida: Florida State University.
- Timm, N. H. (2002). *Applied multivariate analysis*. New York: Springer

LAMPIRAN

Lampiran 1. Output Regresi Logistik

Hosmer and Lemeshow Test

Step	Chi-square	Df	Sig.
1	14,839	6	,022

Classification Table^a

	Observed	Predicted		
		Miskin		Percentage Correct
		tidak miskin	miskin	
Step 1	tidak miskin	5982	0	100,0
	miskin	246	0	,0
	Overall Percentage			96,1

a. The cut value is ,500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	
Step 1 ^a	Jumlah_ruta	1,583	,136	135,330	1	,000	4,870
	JK	,856	,198	18,598	1	,000	2,354
	PDDK	,511	,176	8,443	1	,004	1,667
	Pekerjaan	,344	,143	5,785	1	,016	1,411
	Status_kerja	,535	,156	11,732	1	,001	1,707
	Constant	-4,797	,187	657,662	1	,000	,008

a. Variable(s) entered on step 1: Jumlah_ruta, JK, PDDK, Pekerjaan, Status_kerja.

Lampiran 2. Kombinasi BF, MI, dan MO pada Model MARS

No	BF	MI	MO	GCV	Total Corrected
(1)	(2)	(3)	(4)	(5)	(6)
1	10	1	0	0,037	81,04%
2	10	1	1	0,037	81,04%
3	10	1	2	0,037	81,04%
4	10	1	3	0,037	81,04%
5	10	2	0	0,037	82,93%
6	10	2	1	0,037	82,93%
7	10	2	2	0,037	82,93%
8	10	2	3	0,037	82,93%
9	10	3	0	0,037	82,93%
10	10	3	1	0,037	82,93%
11	10	3	2	0,037	82,93%
12	10	3	3	0,037	82,93%
13	15	1	0	0,037	81,04%
14	15	1	1	0,037	81,04%
15	15	1	2	0,037	81,04%
16	15	1	3	0,037	81,04%
17	15	2	0	0,037	82,93%
18	15	2	1	0,037	82,93%
19	15	2	2	0,037	82,93%
20	15	2	3	0,037	82,93%
21	15	3	0	0,037	82,93%
22	15	3	1	0,037	82,93%
23	15	3	2	0,037	82,93%
24	15	3	3	0,037	82,93%
25	20	1	0	0,037	81,04%
26	20	1	1	0,037	81,04%
27	20	1	2	0,037	81,04%
28	20	1	3	0,037	81,04%
29	20	2	0	0,037	82,93%
30	20	2	1	0,037	82,93%
31	20	2	2	0,037	82,93%
32	20	2	3	0,037	82,93%
33	20	3	0	0,037	82,93%
34	20	3	1	0,037	82,93%
35	20	3	2	0,037	82,93%
36	20	3	3	0,037	82,93%

Lampiran 3. Output Model MARS

Relative Variable Importance

Variable	Importance	-gcv
1 JUMLAH_R	100.000	0.038
5 STATUS_K	41.396	0.037
2 JK	26.976	0.037
4 PEKERJAA	20.011	0.037
3 PDDK	0.000	0.037

ORDINARY LEAST SQUARES RESULTS

N: 6228.000 R-SQUARED: 0.035
 MEAN DEP VAR: 0.039 ADJ R-SQUARED: 0.034
 UNCENTERED R-SQUARED = R-0 SQUARED: 0.073

PARAMETER	ESTIMATE	S.E.	T-RATIO	P-VALUE
Constant	0.153	0.010	14.649	.999201E-15
Basis Function 1	-0.072	0.006	-12.719	.999201E-15
Basis Function 3	-0.037	0.006	-6.489	.933020E-10
Basis Function 5	-0.040	0.009	-4.383	.119176E-04
Basis Function 7	-0.025	0.007	-3.765	.168064E-03

F-STATISTIC = 55.803 S.E. OF REGRESSION = 0.191
 P-VALUE = .999201E-15 RESIDUAL SUM OF SQUARES = 228.102
 [MDF,NDF] = [4, 6223] REGRESSION SUM OF SQUARES = 8.182

The Following Graphics Are Piecewise Linear
 0 curves and 0 surfaces.

Basis Functions

BF1 = (JUMLAH_R = 0);
 BF3 = (STATUS_K = 0);
 BF4 = (STATUS_K = 1);
 BF5 = (JK = 0);
 BF7 = (PEKERJAA = 0) * BF4;

$$Y = 0.153 - 0.072 * BF1 - 0.037 * BF3 - 0.040 * BF5 - 0.025 * BF7;$$

model MISKIN = BF1 BF3 BF5 BF7;

LEARNING SAMPLE CLASSIFICATION TABLE

Actual Class	Predicted Class		Actual Total
	0	1	
0	5054.000	928.000	5982.000
1	135.000	111.000	246.000
Pred. Tot.	5189.000	1039.000	6228.000
Correct	0.845	0.451	
Success Ind.	-0.116	0.412	
Tot. Correct	0.829		
Sensitivity:	0.845	Specificity:	0.451
False Reference:	0.026	False Response:	0.893
Reference = Class 0, Respon = Class 1			