E-ISSN 2527-9378

Jurnal Statistika Industri dan Komputasi
Volume 5, No. 1, Januari 2020, pp. 93-106

PEMODELAN FAKTOR-FAKTOR YANG MEMPENGARUHI INDEKS KETAHANAN PANGAN DENGAN MENGGUNAKAN PENDEKATAN MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS)

Prismakawa Punggodewi [1], Noviana Pratiwi [2]

[1,2] Jurusan Statistika, Fakultas Sains Terapan, IST AKPRIND Yogyakarta
Email: *prismakawa17@gmail.com*

Abstract. Multivariate Adaptive Regression Spline (MARS) is a regression analysis with a nonparametric approach used when the relationship between the response variables to the Predictor variable does not form a specific pattern. MARS is a progression of linear truncated spline regression combined with recursive partitioning regression. The best MARS model is a model that has a combination of base function (BF), maximum interaction (MI), and minimum Observation (MO) which results in the smallest Generalized Cross-Validation (GCV) value. The purpose of this research is to obtain a regression model of the factors that affect the food security index (IKP) so that it can know the contribution of each factor in the value of the IKP. Based on the results of modeling factors that affect the IKP, has been selected model with a combination of BF = 20, MI = 3 and MO = 2 that produces the smallest GCV value of 25.20945 and R2 value of 86.4% indicate the model is good. The selected MARS Model has 12 base functions including the intercept with five variables contained inside the model, with the most important variable is the Village Development Index

Keywords: Food Security Index , Multivariate Adaptive Regression Spline, Generalized Cross Validation

Abstrak Multivariate Adaptive Regression Spline (MARS) merupakan analisis regresi dengan pendekatan nonparametrik yang digunakan ketika hubungan antara variabel respon terhadap variabel prediktor tidak membentuk suatu pola tertentu. MARS merupakan perkembangan dari regresi spline truncated linier yang dikombinasikan dengan recursive partitioning regression. Model MARS terbaik adalah model yang memiliki kombinasi basis function (BF), maximum interaction (MI), dan minimum observation (MO) yang menghasilkan nilai Generalized Cross Validation (GCV) yang paling kecil. Tujuan dari adanya penelitian ini untuk mendapatkan model regresi dari faktor-faktor yang mempengaruhi Indeks Ketahanan Pangan (IKP) sehingga dapat mengetahui kontribusi masing-masing faktor dalam nilai IKP. Berdasarkan hasil pemodelan faktor-faktor yang mempengaruhi IKP, dipilih model kombinasi nilai BF=20, MI=3 dan MO=2 yang menghasilkan nilai GCV terkecil yakni sebesar 25,20945 dan nilai R2 sebesar 86,4% menandakan model tersebut baik. Model MARS yang terpilih memiliki 12 fungsi basis termasuk intercept dengan lima variabel termuat didalamnya, dan variabel dengan tingkat kepentingan tertinggi adalah indeks pembangunan desa.

Kata kunci: ***Indeks Ketahanan Pangan , Multivariate Adaptive Regression Spline, Generalized Cross Validation***

1. Pendahuluan

Sebagai kebutuhan dasar dan hak asasi manusia, pangan memiliki peran penting dalam membentuk sumber daya manusia yang berkualitas demi melaksanakan pembangunan negara. Suatu daerah yang dapat memenuhi kebutuhan pangannya dikatakan dengan daerah tahan pangan. Untuk mengukur ketahanan pangan suatu daerah maka pemerintah membuat suatu indikator untuk melihat capaian ketahanan pangan suatu daerah, indikator tersebut adalah Indeks Ketahanan Pangan. Menurut Indeks Ketahanan Pangan tahun 2018 terdapat penurunan jumlah daerah rawan pangan dari periode sebelumnya sebanyak 110 menjadi 81 kabupaten, sedangkan untuk daerah tahan pangan meningkat dari periode sebelumnya sebanyak 288 menjadi 335 kabupaten. Namun pada tahun 2018 terjadi peningkatan jumlah daerah sangat rentan pangan menurut prioritas penanganannya yakni sebanyak 24 kabupaten lebih banyak dari periode sebelumnya yakni sebanyak 14 kabupaten. Hal ini menunjukkan bahwa disamping mengalami peningkatan ketahanan pangan, di Indonasia masih banyak daerah yang mengalami kerentanan pangan.

Pencapaian ketahanan pangan di tiap-tiap daerah tidak lepas dari adanya faktor yang mempengaruhinya. Untuk itu perlu suatu analisis untuk memodelkan faktor-faktor yang berhubungan dengan ketahanan pangan di Indonesia dilihat melalui Indeks Ketahanan Pangan untuk mengetahui seberapa besar kontribusi faktor-faktor tersebut dalam naik turunnya indeks ketahanan pangan. Analisis Regresi merupakan suatu analisis statistik yang dapat memberikan penjelasan mengenai pola hubungan antara dua variabel atau lebih. Multivariate Adaptive Regression Spline (MARS) adalah salah satu pemodelan regresi nonparametrik yang dapat digunakan pada data ketika hubungan antara variabel respon dan variabel prediktor memiliki sebaran yang tidak jelas, atau dapat dikatakan tidak membentuk suatu pola tertentu. MARS memiliki model yang fleksibel dan digunakan untuk data berdimensi tinggi yakni data dengan variabel prediktor sebanyak dan jumlah observasi sebanyak (Febriyanti, 2012). Model MARS terbaik model yang memiliki kombinasi basis function (BF), maximum interaction (MI), dan minimum observation (MO) yang menghasilkan nilai Generalized Cross Validation (GCV) yang paling kecil. GCV merupakan suatu metode untuk mencari titik knot optimal.

Penelitian ini dilakukan untuk mengetahui [1] karakteristik dari indeks ketahanan pangan dengan faktor-faktor yang mempengaruhinya antara lain kepadatan penduduk, PDRB, persentase partisipasi angkatan kerja, pengeluaran perkapita, dan indeks pembangunan desa. [2] Apakah data Indeks Ketahanan Pangan dan faktor-faktor yang mempengaruhinya memenuhi kriteria pemodelan dengan pendekatan regresi nonparametrik [3] Model MARS yang terbentuk untuk menjelaskan indeks ketahanan pangan dan faktor-faktor yang mempengaruhinya. Tujuan dari penelitian ini adalah untuk mendapatkan model regresi yang menjelaskan faktor-faktor yang mempengaruhi indeks ketahanan pangan sehingga dapat mengetahui kontribusi masing-masing faktor terhadap naik atau turunnya nilai indeks ketahanan pangan.

Penelitian terdahulu yang membahas tentang MARS adalah penelitian Irwan (2017) “Memodelkan Faktor-Faktor Yang Mempengaruhi Gizi Buruk Balita Dengan Metode Multivariate Adaptive Regression Spline (MARS)”. Penelitian ini mengkaji permasalahan gizi buruk balita di provinsi Sulawesi Selatan selama tahun 2015 dengan pendekatan model MARS terbentuk model faktor-faktor yang mempengaruhi gizi buruk dengan variabel prediktor yang paling berpengaruh adalah umur, berat badan, inisiasi menyusui dini, dan pemberian kapsul vitamin A.

2. Metode

2.1 Bahan

Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari publikasi digitan Badan Ketahanan Pangan, Kementrian Pertanian http://bkp.pertanian.go.id dan Badan Pusat Statistik Provinsi se-Indonesia. *Software* yang digunakan untuk menganalisis adalah *software* R. Data terdiri dari variabel respon Y berupa Indeks Ketahanan Pangan dan variabel prediktor X antara lain, X_1 : Kepadatan Penduduk, X_2 : PDRB, X_3 : Persentase Partisipasi Angkatan Kerja, X_4 : Pengeluaran Perkapita, X_5 : Indeks Pembangunan Desa

2.2 Metode

2.2.1 Analisis Regresi

Analisis Regresi merupakan salah satu analisis statistik untuk membangun persamaan yang dapat menggambarkan hubungan antara variabel prediktor dan variabel respon. Terdapat dua cara yang digunakan untuk mengestimasi fungsi regresi yakni dengan pendekatan secara parametrik dan pendekatan secara non-parametrik. Salah satu regresi parametrik yang banyak diaplikasikan pada umumnya adalah regresi linier sederhana dan berganda. Regresi linier sederhana dapat menjelaskan hubungan antara satu variabel respon dengan satu variabel prediktor dan dapat dijelaskan dalam fungsi linier sebagai berikut :

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \tag{1}$$

Regresi linier berganda menjelaskan hubungan antara satu variabel terikat dipengaruhi oleh lebih dari satu variabel bebas. Bentuk umum regresi linier berganda dapat dinyatakan sebagai berikut :

$$Y_i = \beta_0 + \beta_{1i} X_{1i} + \beta_{2i} X_{2i} + ... + \beta_{ki} X_{ki} + \varepsilon_i \tag{2}$$

Y_i = Variabel terikat
X_{ki} = Variabel bebas
β_k = Parameter Regresi
ε_i = Variabel Galat (error)

Persamaan (2) diatas dapat dinyatakan dalam bentuk matriks sebagai berikut :

$$Y = X\beta + \varepsilon \tag{3}$$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}; \; X = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1k} \\ 1 & X_{21} & \cdots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \cdots & X_{nk} \end{bmatrix}; \; \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{k+1} \end{bmatrix}; \; \varepsilon = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Dengan menggunakan metode kuadrat terkecil atau OLS (Ordinary Least Square) dapat dilakukan estimasi parameter model regresi yakni dengan cara meminimumkan jumlah kuadrat errornya. Estimator-estimator regresi linier yang didapat melalui metode OLS dapat dinyatakan dalam bentuk matriks sebagai berikut :

$$\hat{\beta} = (X^T X)^{-1} X'Y \tag{4}$$

2.2.3 Multivariate Adaptive Regression Spline (MARS)

Multivariate Adaptive Regression Splines (MARSplines) merupakan salah satu model regresi nonparametrik, yaitu model yang mengasumsikan bentuk hubungan fungsional antara variabel respon dan prediktor tidak diketahui. MARS adalah implementasi metode statistik yang dipopulerkan oleh Friedman (1991) sebagai pendekatan model regresi multivariate non-parametrik yang fleksibel digunakan untuk pemodelan data berdimensi tinggi yaitu data dengan jumlah variabel prediktor sebesar $3 \leq n \leq 20$, dan jumlah sampel yang diharuskan untuk pendekatan MARS adalah $50 \leq n \leq 1000$.

Beberapa istilah yang perlu diperhatikan dalam metode dan pemodelan MARS adalah sebagai berikut (Febriyanti, 2012) :

1. Knot, merupakan nilai variabel prediktor ketika slope suatu garis regresi mengalami perubahan yang dapat didefinisikan sebagai akhir dari satu segmen sekaligus merupakan awal dari segmen yang lain.
2. Fungsi basis (BF) yaitu selang antar knot yang berurutan. Fungsi basis merupakan fungsi yang dapat menjelaskan hubungan antara variabel prediktor dan variabel respon.
3. Interaction (interaksi) yaitu hasil perkalian silang antar variabel yang saling berkorelasi

Model multivariate adaptive regression splines (MARS) hasil dari pengembangan oleh Friedman (1991) adalah sebagai berikut (Sita & Otok, 2014) :

$$f(x) = a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m} \left[S_{km} \left(x_{v(k,m)i} - t_{(k,m)} \right) \right]_+ \tag{5}$$

dimana :

$$B_{mi}(x)=\prod_{k=1}^{K_m}\left[S_{km}\left(x_{v(k,m)i}-t_{(k,m)}\right)\right]_+ \qquad (6)$$

dengan :

i = 1,2,…,n banyaknya pengamatan
a_0 = Konstanta regresi dari fungs basisi
a_m = Koefisien dari fungsi basis ke-m, m = 1, ..., M
$B_{mi}(x)$ = Fungsi basis ke-m
M = Maksimum fungsi basisi (nonconstant fungsi basis)
K_m = maksimum derajat Interaksi pada fungsi basis ke-m
S_{km} = +1 Jika data berada di sebelah kanan titik knot, Maka

$$\left[+\left(x_{v(k,m)i}-t_{(k,m)}\right)\right]_+=\begin{cases}x_{v(k,m)i}-t_{(k,m)} & ,\ jika\ \ x_{v(k,m)i}>t_{(k,m)}\\ 0 & ,\ sebaliknya\end{cases}$$

-1 Jika data berada di sebelah kiri titik knot, Maka

$$\left[-\left(x_{v(k,m)i}-t_{(k,m)}\right)\right]_+=\begin{cases}x_{v(k,m)i}-t_{(k,m)} & ,\ jika\ \ t_{(k,m)}>x_{v(k,m)i}\\ 0 & ,\ sebaliknya\end{cases}$$

$x_{v(k,m)}$ = Variabel Prediktor ke –v dengan interaksi ke-k dan fungsi basis ke-m
$t_{(k,m)}$ = Nilai knots interaksi ke-k dan fungsi basis ke-m

Model MARS dapat ditulis dalam bentuk matriks sebagai berikut :

$$\underline{Y}=B\,\underline{\hat{a}}+\underline{\varepsilon} \qquad (7)$$

Dengan

$$Y=\left(Y_1,Y_2,...,Y_n\right)^T \text{ dan } \hat{a}=(a_0,a_1,a_2,...,a_n)^T$$

$$B=\begin{bmatrix}1 & \prod_{k=1}^{K_1}\left[S_{k1}\left(x_{v(k,1)1}-t_{k1}\right)\right]_+ & \cdots & \prod_{k=1}^{K_M}\left[S_{kM}\left(x_{v(k,M)1}-t_{kM}\right)\right]_+\\ 1 & \prod_{k=1}^{K_1}\left[S_{k1}\left(x_{v(k,1)2}-t_{k1}\right)\right]_+ & \cdots & \prod_{k=1}^{K_M}\left[S_{kM}\left(x_{v(k,M)2}-t_{kM}\right)\right]_+\\ \vdots & \vdots & \ddots & \vdots\\ 1 & \prod_{k=1}^{K_1}\left[S_{k1}\left(x_{v(k,1)n}-t_{k1}\right)\right]_+ & \cdots & \prod_{k=1}^{K_M}\left[S_{kM}\left(x_{v(k,M)n}-t_{kM}\right)\right]_+\end{bmatrix}$$

$\hat{a}$ dalam model MARS merupakan parameter yang akan diestimasi, melalui penalized least square (PLS). Bentuknya adalah sebagai berikut :

$$PLS=\frac{1}{n}\sum_{i=1}^{n}\left(y_i-f(x)\right)^2+\eta\int_a^b\left(f''(x)\right)^2dx \qquad (8)$$

Sehingga estimasi dengan PLS untuk meminimumkan jumlah kuadrat eror dapat dinyatakan sebagai berikut:

$$a=(B^TB+\eta\,D)^{-1}B^TY \qquad (9)$$

Jika $\eta\to 0$, maka hasil estimasi mendekati hasil metode kuadrat terkecil. Sebaliknya, jika $\eta\to\infty$ maka penalty mendominasi estimasi dan akan menginterpolai titik-titik data (Wibowo, 2013).

Pemilihan model pada MARS dilakukan dalam dua tahap yaitu tahap forward dan tahap backward. Forward dilakukan untuk mendapatkan fungsi dengan jumlah basis maksimum. Pada tahap backward akan dipilih satu fungsi basis dan mengeluarkan basis tersebut jika kontribusi terhadap model kecil. Proses backward akan di lanjutkan hingga tidak ada fungsi basis yang dapat dikeluarkan. Ukuran kontribusi pada tahap backward ditentukan berdasarkan kriteria *Generalized Cross Validation* (GCV) yang didefinisikan sebagai berikut :

$$GCV(M) = \frac{\frac{1}{N}\sum_{i=1}^{N}\left[y_i - f_M(x_i)\right]^2}{\left[1-\left(\tilde{C}(M)/N\right)\right]^2} \tag{10}$$

M = Jumlah fungsi basis
x_i = Variabel prediktor
y_i = Variabel respon
N = Banyaknya pengamatan
$\tilde{C}(M)$ = Fungsi *complexity cost* $\rightarrow \tilde{C}(M) = C(M) + d(M),\ 2 \le d \le 4$
$C(M)$ = $Trace[B(B^T B)^{-1} B^T] + 1$
$f_M(x_i)$ = Nilai estimasi dari variabel respon pada M fungsi basis

4. Hasil dan Pembahasan

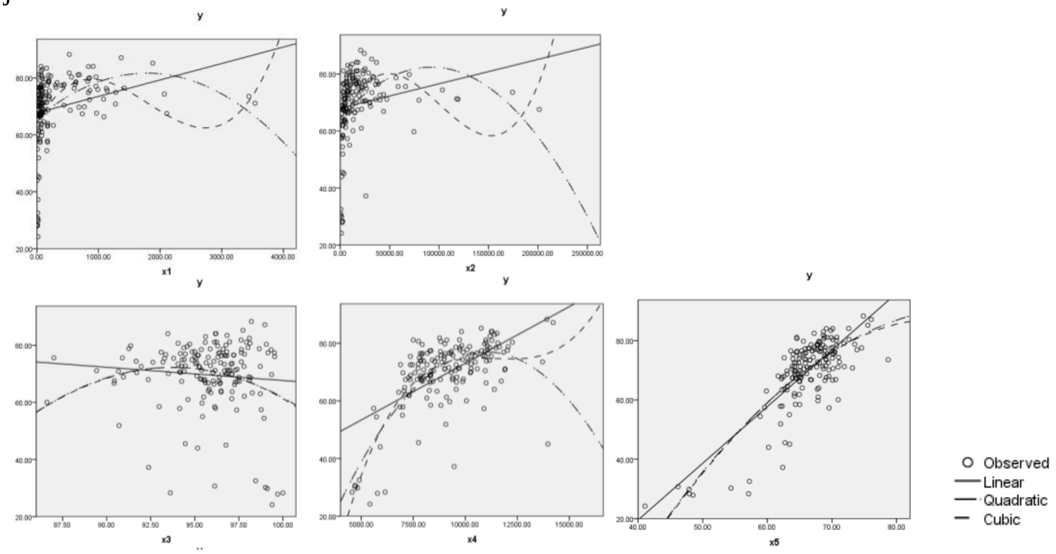
4.1 Analisis Deskriptif

Karakteristik data Indeks Ketahanan Pangan (IKP) yang ingin dilihat melalui analisis deskriptif meliputi nilai rata-rata, nilai maksimum, nilai minimum dan nilai standard deviasi. Analisis deskriptif masing-masing variabel yang dapat dilihat pada Tabel 1 sebagai berikut :

Tabel 1. Analisis Deskriptif

Variabel	N	Minimum	Maximum	Rata-Rata	Std. Deviation
Y	167	24,21	88,3	69,6	12,26
X1	167	1	3.542,0	322,5	539,9
X2	167	171	201.387	18.857	27.268
X3	167	86,61	100,0	95,78	2,354
X4	167	4.554	14.222	9.124	1.862
X5	167	26,10	73,40	59,63	9,59

Untuk mengetahui bagaimana pola sebaran hubungan antara variabel respon dan variabel prediktor maka akan dilakukan estimasi kurva, yang dapat disajikan dalam plot estimasi kurva disajikan dalam Gambar 1 :



Gambar 1. Plot Estimasi Kurva

4.2. Analisis Regresi Linier

Hasil analisis regresi linier berganda diperoleh estimasi parameter model regresi linier berganda untuk memodelkan Indeks Ketahanan Pangan dengan faktor-faktor yang mempengaruhi adalah sebagai berikut :

$$\hat{Y} = -17{,}94 - 1{,}88.10^{-3} X_1 - 4{,}053.10^{-5} X_2 + 2{,}096.10^{-1} X_3 - 1{,}018.10^{-3} X_4 + 0{,}998 X_5$$

1. Pengujian Parameter Serentak

a. Hipotesis

H_0 : $a_0 = a_1 = a_2 = ... = a_M = 0$ (model tidak signifikan)

H_1 : minimal terdapat satu $a_i \neq 0$, i=1,2,…,M (model signifikan)

b. Signifikansi

Tingkat signifikansi menggunakan α (5% atau 0,05) dan $F_{\alpha,k;n-k-1}$

c. Statistik Uji yang digunakan adalah uji F :

F=67,80; p-value = 0,000

d. Daerah Kritis

Daerah Kritis untuk uji ini adalah jika nilai $F_{hitung} > F_{\alpha,k;n-k-1}$ atau *P-value* < *α*, maka H_0 di tolak. Pada tingkat signifikansi dengan α = 5%, p-value berilai kurang dari α, dan F_{hitung} = 67,80 > F_{tabel} = 2,27 sehingga H_0 ditolak.

e. Kesimpulan

Berdasarkan pada hasil daerah kritis untuk uji signifikansi parameter secara simultan menunjukkan H_0 ditolak, sehingga dapat disimpulkan bahwa model signifikan.

2. Pengujian Parameter Parsial

a. Hipotesis

H_0 : $a_i = 0$ (koefisien parameter a_i tidak berpengaruh terhadap model)

H_1 : $a_i \neq 0$, untuk setiap i, dimana i=1,2,…,M (koefisien parameter a_j berpengaruh terhadap model)

b. Signifikansi

Tingkat signifikansi menggunakan α (5% atau 0,05) dan $t_{(\alpha/2,n-k)}$

c. Statistik Uji menggunakan uji T

d. Daerah Kritis

Daerah penolakan pada $t_{hitung} > t_{(\alpha/2,n-k)}$ = 1.97472 atau *P-value* < *α* maka maka H_0 di tolak. Hasil pengujian secara individu dan keputusan disajikan dalam Tabel 3

Tabel 3 Estimasi Parameter Regresi Linier Berganda

Variabel Prediktor	Koefisien	SE Koefisien	T	P	Keputusan
X_1	-116,053	28.1808	-1,251	0.2127	H_0 tidak ditolak
X_2	0.00294450	0.000031	-1,356	0,1771	H_0 tidak ditolak
X_3	-0.00007696	0.260723	0,842	0.4010	H_0 tidak ditolak
X_4	0.49735035	0.000511	2,551	0.0117	H_0 ditolak
X_5	-0.00046518	0.18911	12,333	< 2E-16	H_0 ditolak

e. Kesimpulan

Berdasarkan hasil keputusan dalam daerah kritis terdapat tiga variabel yakni variabel X_1 (Kepadatan Penduduk), X_2 (Produk Domestik Regional Bruto) dan variabel X_3 (Persentase Partisipasi Angkatan Kerja) yang tidak berpengaruh signifikan terhadap model karena memiliki nilai *p-value* > *α*. Sedangkan variabel lainnya, Pengeluaran Perkapita dan Indeks Pembangunan Desa berpengaruh signifikan terhadap Indeks Ketahanan Pangan.

3. Uji Linieritas RESET

a. Hipotesis
   H_0 : Model linier
   H_1 : Model tidak linier
b. Signifikansi
   tingkat signifikan $\alpha = 5\%$
c. Statistik Uji
   F=5,7451
   p-value = 0.0000002746
d. Daerah Kritis
   H_0 ditolak jika F > $F_{(5\%,10,151)}$ = 1,89 atau p-value < α, dalam pengujian linieritas menyatakan F = 5.7451 > $F_{(5\%,10,151)}$ = 1,89 dan p-value = 0.0000002746< α, sehingga H_0 ditolak.
e. Kesimpulan
   Berdasarkan hasil keputusan pada daerah kritis menyatakan bahwa H_0 ditolak, sehingga dapat disimpulakan bahwa data tidak cocok untuk dimodelkan dengan regresi linier.

4. Uji Normalitas

a. Hipotesis
   H_0 : Residual berdistribusi normal
   H_1 : Residual tidak berdistribusi normal
b. Signifikansi
   Tingkat signifikan $\alpha = 5\%$
c. Statistik Uji
   KS = 2,2182
   p-value = 0,3428
d. Daerah Kritis
   Daerah penolakan adalah H_0 ditolak apabila p-value < α. Berdasarkan nilai Kolmogorof-Smirnov menghasilkan *p-value* sebesar 0,3428 lebih dari nilai α, maka H_0 tidak ditolak.
e. Kesimpulan
   Berdasarkan pada hasil keputusan dalam daerah kritis menyatakan H_0 tidak ditolak sehingga dapat disimpulkan bahwa residual berdistribusi normal.

5. Uji Multikolinieritas

a. Hipotesis
   H_0 : $\rho = 0$ (Tidak terindikasi multikolinieritas)
   H_1 : $\rho \neq 0$ (Terindikasi multikolinieritas)
b. Signifikansi
   tingkat signifikan $\alpha = 5\%$
c. Statistik Uji
   Pengujian multikolinieritas dilakukan berdasarkan perolehan nilai VIF.
d. Daerah Kritis
   Daerah penolakan untuk uji hipotesis ini adalah nilai VIF < 10 maka H_0 ditolak. Berikut ini hasil nilai VIF dan keputusan masing-masing variabel ditunjukkan dalam Tabel 4 :

Tabel 4 Nilai VIF

Variabel	Nilai VIF	Keputusan	Variabel	Nilai VIF	Keputusan
X_1	2,195	H_0 ditolak	X_4	1,834	H_0 ditolak
X_2	2,205	H_0 ditolak	X_5	2,003	H_0 ditolak
X_3	1,139	H_0 ditolak			

e. Kesimpulan

Berdasarkan hasil keputusan dalam daerah kritis, apabila H_0 ditolak maka tidak terindikasi adanya multikolinieritas.

6. Uji Autokorelasi

a. Hipotesis
H_0 : tidak ada autokorelasi antar residual
H_1 : ada autokorelasi antar residual

b. Signifikansi
Tingkat signifikansi sebesar $\alpha = 5\%$

c. Statistik Uji
DW = 2,2182
p-value = 0,2208

d. Daerah Kritis
Daerah kritis pada pengujian ini adalah H_0 ditolak jika $d < dL$ atau $d > 4 - dL$ dan H_0 tidak ditolak jika $dU < d < 4 - dU$, sedangkan apabila $dL < d < dU$ atau $4 - dU < d < 4 - dL$ maka hipotesis ini tidak ditarik kesimpulan. Atau H_0 ditolak jika *p-value* < α. Berdasarkan hasil pengujian autokorelasi menyatakan bahwa *p-value* > α, sehingga H_0 tidak ditolak.

e. Kesimpulan
Berdasarkan pada hasil keputusan dalam daerah kritis yang menyatakan bahwa H_0 tidak ditolak, maka dapat disimpulkan bahwa model regresi linier Indeks Ketahanan Pangan tidak terindikasi adanya autokorelasi residual antar pengamatan.

7. Uji Heteroskedastisitas

a. Hipotesis
H_0 : $\sigma_1^2 = \sigma_2^2 = ... = \sigma_n^2 = \sigma^2$ (Varians masing-masing residual bernilai konstan)
H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2;\ i = 1,2,...,n$

b. Signifikansi
Tingkat signifikansi sebesar $\alpha = 5\%$

c. Statistik Uji
BP = 8.6254
p-value = 8.6254

d. Daerah Kritis
Daerah kritis pada pengujian ini adalah H_0 ditolak jika *P-value* < α yang menunjukkan adanya kondisi heteroskedastisitas. Berdasarkan pada Tabel 4.9 menghasilkan *p-value* = 0,003114 kurang dari nilai α, sehingga H_0 ditolak

e. Kesimpulan
Berdasarkan pada hasil keputusan dalam daerah kritis yang menyatakan bahwa H_0 ditolak, maka dapat disimpulkan bahwa model regresi linier Indeks Ketahanan Pangan terindikasi adanya gejala heteroskedastisitas.

Berdasarkan pengujian yang wajib diakukan dalam pemodelan regresi linier berganda, dari ke-empat pengujian asumsi klasik terdapat asumsi yang tidak terpenuhi dan untuk uji linieritas dari model regresi yang terbentuk tidak memiliki sifat linier. Sehingga pemodelan akan dilanjutkan dengan pendekatan regresi non-parametrik dengan metode MARS.

4.2 Estimasi Parameter Model MARS

Pada proses pembentukan model MARS, terdapat tiga hal yang perlu diperhatikan yaitu fungsi basis (BF), maksimum interaksi (MI), dan minimum observasi (MO). Dalam penelitian ini nilai BF, MI dan MO yang digunakan antara lain BF = 10,15,20; MI = 1,2,3; MO = 0,1,2,3,4,5. Pemilihan nilai BF, MI dan MO yang optimal dilakukan dengan cara *trial and error*.

Nilai BF, MI dan MO yang optimal didapatkan dalam penelitian ini adalah BF = 20, MI = 3, MO = 2. Dari kombinasi ini menghasilkan model dengan R^2 sebesar 86,479 % dan nilai GCV sebesar 25,20945. Model ini menghasilkan 13 fungsi basis yang dapat disajikan dalam Tabel 5.

Tabel 5. Fungsi basis Optimal untuk Model MARS Indeks Ketahanan Pangan

Fungsi Basis	Nilai Fungsi Basis	Fungsi Basis	Nilai Fungsi Basis
BF_1	$(10801 - x_2)_+$	BF_8	$(9463 - x_4)_+$
BF_2	$(x_2 - 10801)_+$	BF_9	$(x_5 - 63{,}91)_+$
BF_3	$(x_4 - 9463)_+$	BF_{10}	$(63{,}91 - x_5)_+$
BF_4	$(47{,}22 - x_5)_+$	BF_{11}	$(84 - x_1)_+$
BF_5	$(x_5 - 47{,}22)_+$	BF_{12}	$(x_5 - 61{,}32)_+$
BF_6	$(11101 - x_4)_+$	BF_{13}	$(98{,}1047 - x_3)_+$
BF_7	$(x_4 - 7834)_+$		

Sedangkan model MARS yang terbentuk berdasarkan fungsi basis yang telah terpilih terdapat 11 persamaan fungsi basis dapat dijelaskan dalam Tabel 6.

Tabel 6. Estimasi Parameter Model MARS pada Data Indeks Ketahanan Pangan

Para meter	**Koefisien**	**Fungsi Basis**	**Model MARS**
a_0	65.48328		65.48328
a_1	$3{,}36.10^{-3}$	BF_1	$3{,}36.10^{-3}$ **.** $(10801 - x_2)_+$
a_2	$-7{,}386.10^{-5}$	BF_2	$-7{,}386.10^{-5}$ **.** $(x_2 - 10801)_+$
a_3	$1{,}912.10^{-3}$	BF_3	$1{,}912.10^{-3}$ **.** $(x_4 - 9463)_+$
a_4	-2,681	BF_4	-2,681 **.** $(47{,}22 - x_5)_+$
a_5	0,601	BF_5	0,601 **.** $(x_5 - 47{,}22)_+$
a_6	$-1{,}192.10^{-6}$	BF_1*BF_6	$-1{,}192.10^{-6}$ **.** $(10801 - x_2)_+$ **.** $(11101 - x4)_+$
a_7	$-1{,}259.10^{-6}$	BF_1*BF_7	$-1{,}259.10^{-6}$ **.** $(10801 - x_2)_+$ **.** $(x_4 - 7834)_+$
a_8	$-9{,}457.10^{-4}$	BF_8*BF_9	$9{,}457.10^{-4}$ **.** $(9463 - x_4)_+$ **.** $(x_5 - 63{,}91)_+$
a_9	$3{,}14.10^{-4}$	BF_8*BF_{10}	$3{,}14.10^{-4}$ **.** $(9463 - x_4)_+$ **.** $(63{,}91 - x_5)_+$
a_{10}	$-5{,}746.10^{-5}$	BF_{11}*BF_3*BF_{12}	$-5{,}746.10^{-5}$ **.** $(84 - x_1)_+$ **.** $(x_4 - 9463)_+$ **.** $(x_5 - 61{,}32)_+$
a_{11}	$-2{,}72.10^{-5}$	BF_{13}*BF_8*BF_{10}	$-2{,}72.10^{-5}$ **.** $(98{,}1047 - x_3)_+$ **.** $(9463 - x_4) +$ **.** $(63{,}91 - x_5)_+$

1. Uji Signifikansi Parameter Simultan

a. Hipotesis

H_0 : $a_0 = a_1 = a_2 = ... = a_M = 0$ (model tidak signifikan)

H_1 : minimal terdapat satu $a_m \neq 0$, m=1,2,…,M (model signifikan)

b. Signifikansi

Tingkat signifikansi menggunakan α (5% atau 0,05) dan $F_{\alpha,k;n-k-1}$

c. Statistik Uji yang digunakan adalah uji F :

F = 90,13

p-value = < 2.2e-16

d. Daerah Kritis

Daerah Kritis untuk uji ini adalah jika nilai $F_{hitung} > F_{\alpha,11;155} = 1{,}85$ atau *P-value* < *α*, maka H_0 di tolak. Berdasarkan hasil pada Tabel 7, didapati nilai *P-value* < *α* sehingga H_0 di tolak

e. Kesimpulan

Jika dari hasil daerah kritis menyatakan H_0 ditolak maka dapat disimpulan bahwa minimal terdapat satu a_m yang tidak sama dengan nol, atau model MARS yang terbentuk menunjukkan hubungan yang signifikan antara variabel respon dan variabel prediktor.

2. Uji Signifikansi Parameter Parsial

a. Hipotesis

H_0 : $a_j = 0$ (koefisien parameter a_j tidak berpengaruh terhadap model)

H_1 : $a_i \neq 0$, untuk setiap *i*, dimana *i*=1,2,…,*M* (koefisien parameter a_j berpengaruh terhadap model)

b. Signifikansi

Tingkat signifikansi menggunakan α (5% atau 0,05) dan $t_{(\alpha/2, n-k)}$

c. Statistik Uji

Pengujian signifikansi parameter yang digunakan adalah uji t :

d. Daerah Kritis

Daerah penolakan pada $|t_{hitung}| > t_{(\alpha/2, n-k)} = 1.97472$ atau *P-value* < *α* maka maka H_0 ditolak. Hasil perolehan nilai *t* untuk masing-masing fungsi basis dan keputusan daerah penolakan dapat disajikan dalam Tabel 7

Tabel 7. Hasil Perolehan Nilai t hitung dan P-value Masing-masing Fungsi Basis

Fungsi Basis	t_{hitung}	*P-value*	*Keputusan*	Fungsi Basis	t_{hitung}	*P-value*	*Keputusan*
BF_1	5.680	6.48e-08	H_0 ditolak	BF_1*BF_7	-7.796	8.70e-13	H_0 ditolak
BF_2	-4.635	7.53e-06	H_0 ditolak	BF_8*BF_9	-2.890	0.00441	H_0 ditolak
BF_3	3.906	0.00014	H_0 ditolak	BF_8*BF_{10}	5.544	1.24e-07	H_0 ditolak
BF_4	-9.732	< 2e-16	H_0 ditolak	BF_{11}*BF_3*BF_{12}	-4.652	6.99e-06	H_0 ditolak
BF_5	7.239	1.97e-11	H_0 ditolak	BF_{13}*BF_8*BF_{10}	-3.195	0.00169	H_0 ditolak
BF_1*BF_6	-6.931	1.06e-10	H_0 ditolak				

e. Kesimpulan

Hasil dari daerah kritis menyatakan bahwa, seluruh parameter berpengaruh secara signifikan terhadap model.

3. Uji Normalitas Residual

a. Hipotesis

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

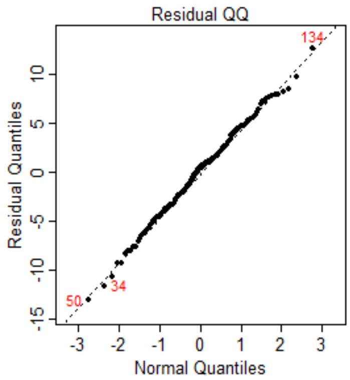
b. Signifikansi

Tingkat signifikansi sebesar $\alpha = 5\%$

c. Statistik Uji

Kolmogorof-Smirnov KS = 0,050751

p-value = 0,7829



Gambar 3. Plot Residual model MARS

Berdasarkan plot residual Gambar 4.7 untuk model MARS, dapat dilihat titik-titik residual berada disekitar garis lurus, hal ini menandakan bahwa residual yang dihasilkan dari model MARS telah berdistribusi normal.

d. Daerah Kritis

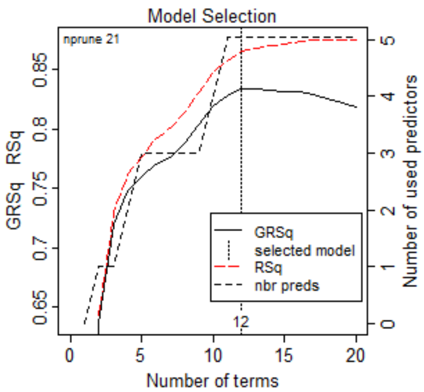
Daerah penolakan adalah H_0 ditolak apabila p-value < α. Berdasarkan nilai Kolmogorof-Smirnov menghasilkan *p-value* sebesar 0.7829 lebih dari nilai α, maka H_0 tidak ditolak.

e. Kesimpulan

Berdasarkan pada hasil keputusan dalam daerah kritis menyatakan H_0 tidak ditolak sehingga dapat disimpulkan bahwa residual berdistribusi normal.

4.3 Interpretasi Model MARS

Koefisien determinasi sebesar 0,864 atau sama dengan 86,4 %, yang mengandung arti bahwa sumbangan pengaruh yang diberikan oleh variabel-variabel yang termuat dalam fungsi basis yang terbentuk dalam model MARS adalah sebesar 86,4 %. Penentuan model terbaik dari pemilihan jumlah variabel dan hasil akhir penyeleksian fungsi basis yang terpilih ke dalam model dapat dilihat pada Gambar 4.



Gambar 4. Penentuan Model Terbaik Indeks Ketahanan Pangan dengan Multivariate Adaptive Regression Spline

Dari Gambar 4, dapat dilihat bahwa penambahan lima variabel merupakan penambahan paling optimal karena mendapatkan R-sq yang paling maksimum, namun untuk pemilihan banyaknya persamaan fungsi basis terhenti sampai penambahan persamaan ke 12 (11 basis fungsi dan 1 *intercept*-nya) karena penambahan persamaan fungsi basis diatas 12 mengakibatkan GR-square mengalami penurunan. Interpretasi masing-masing fungsi basis dapat dijelaskan dalam Tabel 10.

Tabel 10. Interpretasi Masing-masing fungsi basis

Fungsi Basis	Koefisien	Kepadatan Penduduk	PDRB	Persentase Angkatan Kerja	Pengeluaran Perkapita	Indeks Pembangunan Desa
0	65,4832					
1	$3,36.10^{-3}$		$x_2 < 10801$			
2	$-7,386.10^{-5}$		$x_2 > 10801$			
3	$1,912.10^{-3}$				$x_4 > 9463$	
4	-2,681					$x_5 < 47,22$
5	0,601					$x_5 > 47,22$
6	$-1,192.10^{-6}$		$x_2 < 10801$		$x_4 < 11101$	
7	$-1,259.10^{-6}$		$x_2 < 10801$		$x_4 > 7834$	
8	$-9,457.10^{-4}$				$x_4 < 9463$	$x_5 > 63,91$
9	$3,14.10^{-4}$				$x_4 < 9463$	$x_5 < 63,91$
10	$-5,746.10^{-5}$	$x_1 < 84$			$x_4 > 9463$	$x_5 > 61,32$
11	$-2,72.10^{-5}$			$x_3 < 98,10$	$x_4 < 9463$	$x_5 < 63,91$

Berikut ini adalah interpretasi fungsi basis model MARS :

a. Fungsi Basis 1

Koefisien = $3,36.10^{-3}$

$$Fungsi\ Basis\ 1 = \begin{cases} (10801 - x_2) & ; X_2 < 10801 \\ 0 & ; X_2 \geq 10801 \end{cases}$$

Artinya setiap kenaikan satu satuan Fungsi Basis 1 akan meningkat nilai Indeks Ketahanan Pangan sebesar $3,36.10^{-3}$ jika fungsi basis yang lain dianggap konstan, hal ini terjadi pada daerah yang memiliki nilai PDRB kurang dari Rp. 10,801 Miliar.

b. Fungsi Basis 6

Koefisien = $-1,192.10^{-6}$

$$Fungsi\ Basis\ 6 = \begin{cases} (10801 - x_2)_+ (11101 - x_4)_+ & ; X_2 < 10801; X_4 < 11101 \\ 0 & ; X_2 \geq 10801; X_4 \geq 11101 \end{cases}$$ Artiny

a setiap kenaikan satu satuan Fungsi Basis 6 akan menurunkan nilai Indeks Ketahanan Pangan sebesar $1,192.10^{-6}$, jika fungsi basis yang lain dianggap konstan, hal ini terjadi pada daerah dengan nlai PDRB kurang dari Rp. 10,801 miliar dan pengeluaran perkapita kurang dari Rp. 11.101 ribu.

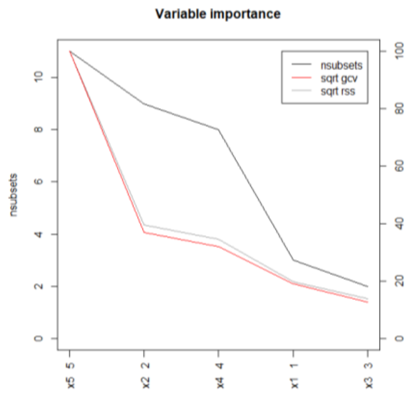
c. Fungsi Basis 10

Koefisien = $-5,746.10^{-5}$

$$Fungsi\ Basis\ 10 = \begin{cases} (84 - x_1)_+ \times (x_4 - 9463)_+ \times (x5 - 61,32)_+ \\ 0 \end{cases}$$

Artinya setiap kenaikan satu satuan Fungsi Basis 10 akan menurunkan nilai Indeks Ketahanan Pangan sebesar $5,746.10^{-5}$ jika fungsi basis yang lain dianggap konstan, hal ini terjadi pada daerah dengan nilai kepadatan penduduk kurang dari dari 84 penduduk/km^3, pengeluaran perkapita lebih dari Rp. 9463 ribu dan indeks pembangunan desa lebih dari 61,32.

Dari lima variabel prediktor yang diduga berpengaruh terhadap Indeks Ketahanan Pangan, semua variabel prediktor dinyatakan berpengaruh terhadap Indeks Ketahanan Pangan, dengan tingkat kepentingan variabel dapat disajikan dalam Gambar 5 :



Gambar 5. Tingkat Kepentingan Masing-masing Variabel Prediktor dalam Model MARS untuk memodelkan data Indeks Ketahanan Pangan

Persentase tingkat kepentingan variabel berdasarkan pada Gambar 5 dapat dijelaskan dalam Tabel 11.

Tabel 11. Tingkat Kepentingan Variabel

No	Variabel	Persentase
1	Indeks Pembangunan Desa	100 %
2	Produk Domestik Regional Bruto	37 %
3	Pengeluara Perkapita	32 %
4	Kepadatan Penduduk	19,1 %
5	Persentase Partisipasi Angkatan Kerja	12,8

5. Kesimpulan

Berdasarkan hasil analisis dan pembahasan, diperoleh kesimpulan sebagai berikut :

1. Analisis deskriptif untuk data Indeks Ketahanan Pangan adalah :
   a. Rata-rata indeks ketahanan pangan kabupaten di Indonesia adalah sebesar 69,603. Dengan standar deviasi 12,261 nilai maksimum sebesar 88,3 dan minimum sebesar 24,21
   b. Rata-rata kepadatan penduduk kabupaten di Indonesia adalah sebesar 322,5. Dengan standar deviasi 539,9 nilai maksimum sebesar 3452 dan minimum sebesar 1.
   c. Rata-rata produk domestik regional bruto kabupaten di Indonesia adalah sebesar 18857. Dengan standar deviasi 27268 nilai maksimum sebesar 201387 dan minimum sebesar 171.
   d. Rata-rata persentase partisipasi angkatan kerja di kabupaten se-Indonesia adalah sebesar 95,788. Dengan standar deviasi 2,354 nilai maksimum sebesar 100 dan minimum sebesar 86,61.
   e. Rata-rata indeks pembangunan desa di kabupaten se- Indonesia adalah sebesar 59,65. Dengan standar deviasi 9,59 nilai maksimum sebesar 73,40 dan minimum sebesar 26,10
2. Data Indeks Ketahanan Pangan dan faktor-faktor yang mempengaruhinya memenuhi kriterian untuk analisis regresi non-parametrik, karena analisis secara regresi linier tidak memenuhi asumsi klasik sehingga tidak mendapatkan estimator linier terbaik dan tidak bias.
3. Model yang terpilih menjadi model terbaik adalah model dengan kombinasi nilai BF = 20, MI = 3, MO = 5 dengan nilai GCV sebesar 25,2095 menghasilkan koefisien determinasi sebesar 86,4 % dan menandakan bahwa model tersebut baik. Model MARS yang terbentuk memiliki 12 fungsi basis termasuk *intercept* dengan lima variabel termuat didalamnya. Berdasarkan pada model MARS, variabel yang memiliki tingkat kepentingan paling tinggi adalah Indeks Pembangunan Desa.

Ucapan Terima Kasih

Dalam penyusunan tulisan ini, banyak pihak yang telah memberikan dukungan selama penyelesaian penelitian ini. Oleh karena itu, pada kesempatan ini peneliti ingin menyampaikan terimakasih kepada Institut Sains & Teknologi Yogyakarta terutama pada Bapak/Ibu dosen Jurusan Statistika atas arahan dan bimbingannya.

Daftar Pustaka

Ampulembang AP, (2017), *Pengembangan Model Regresi Nonparametrik Birespon Kontinu Menggunakan Metode MARS*, disertasi, ITS, Surabaya.

Draper N & Smith H, (1998), *Applied Regression Analysis*, *Third Editio*, Wiley-Interscience, New York.

Eubank R L, (1988), *Spline Smoothing and Nonparametric Regression,* Marcel Dekker, New York.

Febriyanti A, (2012), Penerapan Metode Multivariate Adaptive Regression Spline (MARS) Untuk Mengidentifikasi Komponen Yang Berpengaruh Terhadap Peringkat Akreditasi Sekolah, *Jurnal Matematika FMIPA UNAND*, Vol. 2 No. 2 Hal. 44, Padang.

Moulart R, (2014), On The Use of a Penalized Least Square Method to Process Kinematic Full-Filed Measurements, *Measurement Science and Technology*, England.

Noeryanti, (2012), *Metode Statistika I,* Akprind Press, Yogyakarta.

Otok WB, (2007), Estimasi Spline Dan MARS Menggunakan Kuadrat Terkecil. *Jurnal Matematika, Statistika & Komputasi,* Makasar.

Permatasari DL, (2016), Pemodelan Ketahanan Pangan di Indonesia dengan Pendekatan Regresi Probit Ordinal, *Jurnal Sains dan Seni ITS*, 151-156, Surabaya.

Sayuti A, (2013), Generalized Cross Validation dalam Regresi Smoothing Spline, *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster)*, 191-196, Pontianak.

Setyawan Y, Noeryanti, & Suryowati K., (2018), *Statistika Dasar Dilengkapi dengan Software R,* Akprind Press,Yogyakarta.

Silaban D, (2018), *Pemodelan Faktor-Faktor yang Mempengaruhi Tingkat Kemiskinan di Sumatera Utara dengan Menggunakan Regresi Berganda dan Regresi Nonparametrik Spline,* skripsi, IST Akprind, Yogyakarta.

Sita, & Otok BW., (2014), Pendekatan Multivariate Adaptive Regression SPLINES (MARS) pada Pemodelan Penduduk Miskin di Indonesia Tahun 2008-2012, *Prosiding Seminar Nasional Matematika, Universitas Jember*, Jember.

Wastiti D, (2016), Analisis Regresi Nonparametrik Spline Multivariabel Untuk Pemodelan Indikator Kemiskinan di Indonesia, *E-Jurnal Matematika Vol.5*, 111-116, Bali.