ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 6, Nomor 1, Tahun 2017, Halaman 11-20
Online di: http://ejournal-s1.undip.ac.id/index.php/gaussian

PEMODELAN KASUS KEMISKINAN DI JAWA TENGAH MENGGUNAKAN REGRESI NONPARAMETRIK METODE B-SPLINE

Anisa Septi Rahmawati[1], Dwi Ispriyanti[2], Budi Warsito[3]
[1]Mahasiswa Departemen Statistika FSM Universitas Diponegoro
[2,3]Dosen Departemen Statistika FSM Universitas Diponegoro
anisasepti27@gmail.com

ABSTRACT

Poverty is one of the diseases in the economy, so it must be cured or at least reduced. According to BPS (2016), poor people are people who have an average expenditure per capita per month below the poverty line. The poverty line in Central Java in 2016 amounted to Rp 317 348, - per capita per month. In 2016, the average level of poverty in the Java Island, Central Java province placed as the second highest after DIY. Many factors are thought to affect the level of poverty. In this study, the predictor variables used are the rate of economic growth (X_1), unemployment rate (X_2), and education level above high school to (X_3). This study aims to obtain a model of the relationship between the factors that affect poverty on the percentage of poor and calculate the predictions. The method used is B-spline nonparametric regression. Nonparametric approach are used if the function of previous data is unknown. The best B-spline model depends on the determination of the optimal knots point having a minimum *Generalized Cross Validation* (GCV). In this study, the best B-spline model obtained when the order of X_1 is 2, the order of X_2 is 2, and the order of X_3 is 2. The knots obtained in X_1 at the point 4,51273, X_2 at the point 3,60626, and X_3 at point 11,4129 and 16,2481 with GCV value of 9,79353.

Keywords: Poverty, Nonparametric Regression, B-Spline, *Generalized Cross Validation*

1. PENDAHULUAN

Dalam pembangunan nasional salah satu sasarannya adalah menurunkan tingkat kemiskinan. Kemiskinan merupakan salah satu penyakit dalam ekonomi, sehingga harus disembuhkan atau paling tidak dikurangi. Menurut BPS (2016), kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan dibawah garis kemiskinan. Jumlah penduduk miskin di Jawa Tengah pada tahun 2016 sebesar 4,507 juta orang atau 13,27 persen. Rata-rata tingkat kemiskinan di Jawa Tengah masih yang tertinggi kedua setelah DIY di pulau Jawa.

Beberapa penelitian tentang kemiskinan telah dilakukan, yaitu oleh Merdekawati (2013), melakukan pemodelan menggunakan regresi spline *truncated* multivariabel memperoleh hasil bahwa pertumbuhan ekonomi, tingkat pengangguran terbuka dan tingkat pendidikan berpengaruh terhadap kemiskinan. Prastyo (2010), menganalisis faktor-faktor yang mempengaruhi tingkat kemiskinan menggunakan analisis data panel dan diperoleh hasil bahwa pertumbuhan ekonomi, pendidikan dan pengangguran berpengaruh signifikan terhadap kemiskinan di Jawa Tengah. Wulandari (2014), dengan menggunakan regresi spline birespon memperoleh hasil bahwa pertumbuhan ekonomi dan tingkat pengangguran terbuka berpengaruh terhadap kemiskinan. Berdasarkan penelitian tersebut, peneliti ingin menganalisis faktor-faktor yang mempengaruhi persentase penduduk miskin yaitu pertumbuhan ekonomi, tingkat pengangguran terbuka dan tingkat pendidikan menggunakan regresi nonparametrik metode B-spline.

Hubungan antara faktor-faktor yang mempengaruhi kemiskinan terhadap persentase penduduk miskin dapat dianalisis menggunakan analisis regresi. Terdapat dua pendekatan estimasi model dalam analisis regresi yaitu regresi parametrik dan regresi nonparametrik. Pendekatan parametrik digunakan ketika bentuk fungsi data yang diperoleh menunjukan suatu pola hubungan yang mudah digambarkan dengan fungsi tertentu seperti linier, kuadratik, dan sebagainya. Namun jika fungsi data yang diperoleh tidak menunjukan pola hubungan yang mudah digambarkan dengan fungsi tertentu maka estimasi fungsi regresi dilakukan dengan menggunakan pendekatan nonparametrik.

Regresi nonparametrik memiliki fleksibilitas yang tinggi karena bentuk estimasi kurva regresinya dapat menyesuaikan datanya tanpa dipengaruhi oleh faktor subyektifitas peneliti (Eubank, 1999). Salah satu pendekatan dalam regresi nonparametrik adalah estimator spline. Dalam pendekatan spline memiliki suatu basis fungsi yang biasanya dipakai yaitu *truncated power basis* dan basis B-spline (Eubank, 1999). Spline dengan *truncated power basis* memiliki kelemahan saat orde spline tinggi, knot yang banyak dan knot yang terlalu dekat akan membentuk matriks persamaan normal yang hampir singular, sehingga persamaan normal sulit diselesaikan. Basis lain yang dapat mengatasi kelemahan ini adalah basis B-spline (Eubank, 1999).

Ada kriteria yang harus diperhatikan dalam membentuk model regresi B-spline yaitu menentukan orde untuk model, banyaknya knot, dan lokasi knot. Dengan menggunakan regresi nonparametik B-spline dapat diperoleh model hubungan antara persentase penduduk miskin terhadap faktor-faktor yang mempengaruhinya. Model terbaik diperoleh dengan meminimumkan *Generalized Cross Validation* (GCV).

2. TINJAUAN PUSTAKA

2.1 Pengertian Kemiskinan, Pertumbuhan Ekonomi, Pengangguran, Pendidikan

Menurut BPS (2016), kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan dibawah garis kemiskinan. Garis Kemiskinan (GK) merupakan penjumlahan dari Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Non Makanan (GKNM). Garis Kemiskinan di Jawa Tengah 2016 sebesar Rp 317.348,- per kapita per bulan.

Faktor-faktor yang mempengaruhi kemiskinan salah satunya adalah pertumbuhan ekonomi, tingkat pengangguran terbuka, dan kemiskinan. Untuk mengukur pertumbuhan ekonomi di suatu wilayah perekonomian dalam selang waktu tertentu, maka dilihat dari besarnya nilai laju pertumbuhan ekonomi dengan menggunakan data Produk Domestik Regional Bruto (PDRB). Laju pertumbuhan PDRB dapat diperoleh dari perhitungan PDRB atas dasar harga konstan, dengan rumus sebagai berikut:

$$Laju\ Pertumbuhan\ PDRB = \frac{PDRB_t - PDRB_{t-1}}{PDRB_{t-1}} x100\%$$

Menurut BPS (2016), pengangguran terbuka adalah penduduk yang telah masuk dalam angkatan kerja tetapi tidak memiliki pekerjaan dan sedang mencari pekerjaan, mempersiapkan usaha, serta sudah memiliki pekerjaan tetapi belum mulai bekerja. Tingkat pengangguran terbuka adalah persentase jumlah pengangguran terbuka terhadap jumlah angkatan kerja.

$$TPT = \frac{Jumlah\ pengangguran\ terbuka}{jumlah\ angkatan\ kerja} x100\%$$

Besarnya tingkat pengangguran terbuka mempunyai implikasi sosial yang luas karena semakin tinggi tingkat pengangguran terbuka maka semakin besar potensi kerawanan sosial yang ditimbulkan contohnya kemiskinan.

Pendidikan merupakan salah satu bentuk modal manusia (*human capital*) yang menunjukan kualitas Sumber Daya Manusia (SDM). Melalui investasi pendidikan yang tinggi mampu meningkatkan kualitas SDM yang diperlihatkan oleh meningkatnya pengetahuan dan keterampilan seseorang. Dalam penelitian Winarendra (2014), yang termasuk dalam pendidikan tinggi adalah jenjang pendidikan SMA ke atas, seperti diploma, sarjana, magister, spesialis dan doktor. Meningkatkan tingkat pendidikan jelas merupakan sebuah prioritas untuk memperbaiki standar hidup dan mengurangi kemiskinan.

2.2 Analisis Regresi

Analisis regresi berkenaan dengan studi ketergantungan satu variable respon pada satu atau lebih variabel prediktor, dengan maksud menaksir atau meramalkan nilai rata-rata hitung variabel respon, dipandang dari segi nilai yang diketahui atau tetap (Gujarati, 2004). Hubungan antara variable respon dan variabel prediktor untuk n buah pengamatan, dalam suatu persamaan regresi secara umum dapat ditulis sebagai berikut :

$$y_i = f(x_i, \beta) + \varepsilon_i \text{ dimana } i = 1,2, \ldots, n \qquad (1)$$

dengan ε_i adalah residual pada pengamatan ke-*i* yang merupakan variabel random independen dengan mean 0 dan varians konstan σ^2 dan $f(x_i, \beta)$ merupakan fungsi regresi yang diketahui bentuknya dan mempunyai p buah parameter β.

Pendekatan menggunakan nonparametrik dilakukan jika asumsi bentuk kurva $f(x_i)$ tidak diketahui sebelumnya. Kurva tersebut hanya diasumsikan *smooth* (mulus), artinya termuat di dalam suatu ruang fungsi tertentu. Regresi nonparametrik memiliki fleksibilitas yang tinggi karena bentuk estimasi kurva regresinya dapat menyesuaikan datanya tanpa dipengaruhi oleh faktor subyektifitas peneliti (Eubank, 1999). Model regresi nonparametrik dengan pengamatan (x_i, y_i) adalah:

$$y_i = f(x_i) + \varepsilon_i \text{ dimana } i = 1,2, \ldots, n \qquad (2)$$

dengan $f(x_i)$ merupakan fungsi regresi yang tidak diketahui bentuknya, sementara ε_i adalah residual pada pengamatan ke-*i* yang merupakan variabel random independen dengan mean 0 dan varian konstan σ^2.

2.3 Spline dan Basis B-spline

Spline merupakan salah satu jenis piecewise polinomial, yaitu polinomial yang memiliki sifat tersegmen. Sifat tersegmen ini memberikan fleksibilitas yang lebih dari polinomial biasa sehingga memungkinkan untuk menyesuaikan diri secara lebih efektif terhadap karakteristik lokal suatu fungsi data. Spline mempunyai keunggulan dalam mengatasi pola data yang menunjukan naik atau turun yang tajam dengan bantuan titik-titik knot. Pendekatan spline mempunyai suatu basis fungsi yang biasa dipakai antara lain *truncated power basis* dan basis B-spline. Spline dengan *truncated power basis* mempunyai kelemahan saat orde tinggi, knot yang banyak dan knot yang terlalu dekat akan membentuk matriks dalam persamaan normal untuk menghitung parameter β yang hampir singular sehingga persamaan normal sulit diselesaikan (Budiantara dkk, 2006).

Untuk mengatasi permasalahan pada matriks persamaan normal yang singular digunakan pendekatan spline menggunakan basis fungsi yaitu basis B-spline (Eubank, 1999). Apabila model regresi pada persamaan (2) didekati dengan fungsi B-spline berorde *m* dengan *k* knot, dapat ditulis sebagai berikut :

$$y_i = \sum_{j=1}^{m+k} \beta_j B_{j-m,m}(x_i) + \varepsilon_i \qquad , i = 1,2,\dots,n \qquad (3)$$

dengan $B_{j-m,m}(x)$ merupakan basis B-spline dan β_j merupakan parameter regresi untuk B-spline.

Untuk membangun fungsi B-spline yang berorde *m* dengan *k* titik knot $\xi_1, \xi_2, \dots, \xi_k$ dimana $a < \xi_1 < \cdots < \xi_k < b$, terlebih dahulu didefinisikan knot tambahan sebanyak 2m, yaitu $\xi_{-(m-1)}, \dots, \xi_{-1}, \xi_0, \xi_{k+1}, \dots, \xi_{k+m}$, dengan $\xi_{-(m-1)} = \cdots = \xi_0 = a$ dan $\xi_{k+1} = \cdots = \xi_{k+m} = b$. Biasanya a diambil dari nilai minimum x dan b diambil dari nilai maksimum x (Budiantara dkk, 2006).

Menurut Eubank (1999), basis fungsi B-spline pada orde m dengan titik-titik knot pada $\xi_1, \dots, \xi_k$ dapat didefinisikan secara rekursif sebagai berikut:

$$B_{l,m}(x) = \frac{x-\xi_l}{\xi_{l+m-1}-\xi_l} B_{l,m-1}(x) + \frac{\xi_{l+m}-x}{\xi_{l+m}-\xi_{l+1}} B_{l+1,m-1}(x) \qquad (4)$$

dengan $l = -(m-1), \dots, k$, dan

$$B_{l,1}(x) = \begin{cases} 1, & x \in [\xi_l, \xi_{l+1}] \\ 0, & untuk\ lainnya \end{cases} \qquad (5)$$

2.3.1 Macam-macam Basis B-spline

Berdasarkan persamaan (3), macam-macam basis fungsi B-spline dibedakan berdasarkan jumlah orde *m*, yaitu:

a. Orde m=2 memberikan basis fungsi B-spline linier, yang memiliki fungsi sebagai berikut :

$$B_{l,2}(x) = \frac{x-\xi_l}{\xi_{l+1}-\xi_l} B_{l,1}(x) + \frac{\xi_{l+2}-x}{\xi_{l+2}-\xi_{l+1}} B_{l+1,1}(x) \qquad \text{dengan } l = -1, \dots, k$$

b. Orde m=3 memberikan basis fungsi B-spline kuadratik, yang memiliki fungsi sebagai berikut:

$$B_{l,3}(x) = \frac{x-\xi_l}{\xi_{l+2}-\xi_l} B_{l,2}(x) + \frac{\xi_{l+3}-x}{\xi_{l+3}-\xi_{l+1}} B_{l+1,2}(x) \qquad \text{dengan } l = -2, \dots, k$$

c. Orde m=4 memberikan basis fungsi B-spline kubik, yang memiliki fungsi sebagai berikut:

$$B_{l,4}(x) = \frac{x-\xi_l}{\xi_{l+3}-\xi_l} B_{l,3}(x) + \frac{\xi_{l+2}-x}{\xi_{l+4}-\xi_{l+1}} B_{l+1,3}(x) \qquad \text{dengan } l = -3, \dots, k$$

2.3.2 Estimasi Parameter dalam Model B-spline

Kurva regresi f pada persamaan (2) jika didekati dengan fungsi B-spline berorde *m* dengan *k* titik knot untuk $\lambda = \{\xi_1, \dots, \xi_k\}$ maka menjadi :

$$f_\lambda(x_i) = \sum_{j=1}^{m+k} \beta_{\lambda j} B_{\lambda j-m,m}(x_i)$$

sehingga model B-spline adalah :

$$y_i = \sum_{j=1}^{m+k} \beta_{\lambda j} B_{\lambda j-m,m}(x_i) + \varepsilon_i \qquad (6)$$

Model regresi (6) dapat ditulis menjadi :

$$y_i = \beta_{\lambda 1} B_{\lambda 1-m,m}(x_i) + \beta_{\lambda 2} B_{\lambda 2-m,m}(x_i) + \cdots + \beta_{\lambda(m+k)} B_{\lambda k,m}(x_i) + \varepsilon_i$$

Model B-spline diatas apabila dibuat dalam bentuk matriks maka menjadi:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} B_{\lambda 1-m,m}(x_1) & B_{\lambda 2-m,m}(x_1) & \dots & B_{\lambda k,m}(x_1) \\ B_{\lambda 1-m,m}(x_2) & B_{\lambda 2-m,m}(x_2) & \cdots & B_{\lambda k,m}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_{\lambda 1-m,m}(x_n) & B_{\lambda 2-m,m}(x_n) & \dots & B_{\lambda k,m}(x_n) \end{bmatrix} \begin{bmatrix} \beta_{\lambda 1} \\ \beta_{\lambda 2} \\ \vdots \\ \beta_{\lambda(m+k)} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dapat ditulis menjadi :

$$\boldsymbol{y} = \boldsymbol{B}_{\boldsymbol{\lambda}} \boldsymbol{\beta}_{\boldsymbol{\lambda}} + \boldsymbol{\varepsilon}$$

Estimasi parameter $\boldsymbol{\beta_\lambda}$ diperoleh dengan menggunakan metode *least squares spline*. Estimator $\boldsymbol{\hat{\beta}_\lambda}$ didapat dengan meminimumkan Jumlahan Kuadrat Error atau *Residual Sum of Squares (RSS)*, sehingga didapat :

$$\boldsymbol{\hat{\beta}_\lambda} = \left(\boldsymbol{B_\lambda^T B_\lambda}\right)^{-\boldsymbol{1}} \boldsymbol{B_\lambda^T y}$$

dengan $\boldsymbol{\hat{\beta}_\lambda} = \left(\hat{\beta}_{\lambda 1}\ \ \hat{\beta}_{\lambda 2}\ \ \ldots\ \ \hat{\beta}_{\lambda(m+k)}\right)^T$

Estimasi model untuk fungsi B-spline pada regresi nonparametrik adalah:

$$\begin{aligned}\boldsymbol{\hat{y}} &= \boldsymbol{B_\lambda \hat{\beta}_\lambda} \\ &= \boldsymbol{B_\lambda}\left(\left(\boldsymbol{B_\lambda^T B_\lambda}\right)^{-\boldsymbol{1}} \boldsymbol{B_\lambda^T y}\right) \\ &= \boldsymbol{B_\lambda}\left(\boldsymbol{B_\lambda^T B_\lambda}\right)^{-\boldsymbol{1}} \boldsymbol{B_\lambda^T y} \\ &= \boldsymbol{S_\lambda y}\end{aligned}$$

dengan matriks $\boldsymbol{S_\lambda} = \boldsymbol{B_\lambda}\left(\boldsymbol{B_\lambda^T B_\lambda}\right)^{-\boldsymbol{1}} \boldsymbol{B_\lambda^T}$ simetris dan definit positif (Eubank, 1999).

Estimasi model B-spline dapat juga ditulis sebagai berikut:

$$\hat{y} = \sum_{j=1}^{m+k} \hat{\beta}_{\lambda j} B_{\lambda j-m,m}(x) \tag{7}$$

Persamaan (7) dapat ditulis menjadi:

$$\hat{y} = \hat{\beta}_{\lambda 1} B_{\lambda 1-m,m}(x) + \hat{\beta}_{\lambda 2} B_{\lambda 2-m,m}(x) + \cdots + \hat{\beta}_{\lambda(m+k)} B_{\lambda k,m}(x)$$

2.4 Pemilihan Model B-spline Terbaik

Untuk memperoleh model B-spline terbaik diperluhkan lokasi dan banyaknya knot yang optimal. Salah satu metode pemilihan titik knot yang optimal adalah *Generalized Cross Validation* (GCV). Menurut Eubank (1999), metode GCV dapat dituliskan sebagai berikut:

$$GCV(\lambda) = \frac{MSE(\lambda)}{(n^{-1} trace[I - S_\lambda])^2}$$

dengan n adalah banyaknya pengamatan, I merupakan matriks identitas nxn, $MSE(\lambda) = n^{-1}\sum_{i=1}^{n}(y_i - \hat{f}_\lambda)^2$ dan $\boldsymbol{S_\lambda} = \boldsymbol{B_\lambda}\left(\boldsymbol{B_\lambda^T B_\lambda}\right)^{-\boldsymbol{1}} \boldsymbol{B_\lambda^T}$. Pemilihan model terbaik dilakukan dengan cara membandingkan nilai GCV dari masing-masing orde dan titik knot yang memiliki nilai GCV minimum.

2.5 Koefisien Determinasi ($\boldsymbol{R^2}$)

Koefisien determinasi digunakan untuk menunjukan seberapa besar persentase keragaman dalam variabel respon yang dijelaskan oleh variabel prediktor. Nilai dari R^2 adalah $0 \le R^2 \le 1$, jika R^2 bernilai 1 berarti suatu kecocokan sempurna, sedangkan jika R^2 bernilai 0 berarti tidak ada hubungan antara variabel tak bebas dengan variabel bebas (Gujarati, 2004). Koefisien determinasi dapat dihitung dengan rumus berikut:

$$R^2 = \frac{JKR}{JKT} = 1 - \frac{JKG}{JKT} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}$$

3. METODOLOGI PENELITIAN

Data yang digunakan pada penelitian ini merupakan data sekunder yang diperoleh dari buku Data dan Informasi Kemiskinan Kabupaten/Kota tahun 2015, Produk Domestik Regional Bruto Kota di Indonesia 2011-2015, Keadaan Angkatan Kerja Jawa Tengah Agustus 2015, dan Statistik Sosial dan Kependudukan Jawa Tengah hasil Susenas 2015. Unit pengamatan yang digunakan adalah 35 kabupaten/kota di Provinsi Jawa Tengah.

Variabel yang digunakan adalah persentase penduduk miskin sebagai variabel respon (Y), laju pertumbuhan ekonomi sebagai variabel prediktor (X_1), tingkat pengangguran terbuka sebagai variabel prediktor (X_2) dan tingkat pendidikan terakhir SMA ke atas sebagai variabel prediktor (X_3).

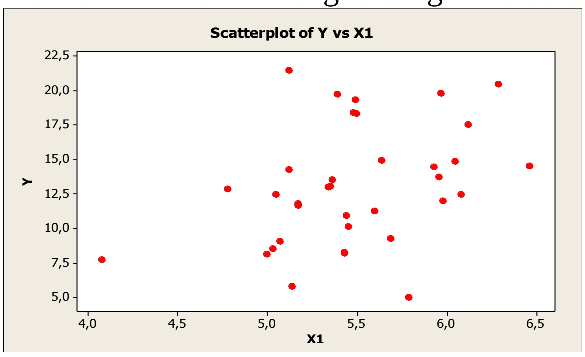
Langkah-langkah analisis yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Memasukan data variabel respon dan variabel prediktor.
2. Membuat *scatterplot* antara variabel respon (y_i) dengan masing-masing variabel prediktor (x_{pi}).
3. Meregresikan variabel menggunakan regresi nonparametrik B-spline dengan menentukan banyaknya titik knot yang diinginkan dari orde 2, orde 3 dan orde 4.
4. Mengkombinasikan beberapa orde dan titik knot.
5. Menghitung nilai GCV untuk setiap kombinasi orde dan titik knot.
6. Menentukan titik knot optimal yang dilihat dari nilai GCV minimum pada masing-masing kombinasi orde dan titik knot.
7. Menentukan model B-spline terbaik berdasarkan titik knot optimal.
8. Menghitung prediksi persentase penduduk miskin dari model yang diperoleh.

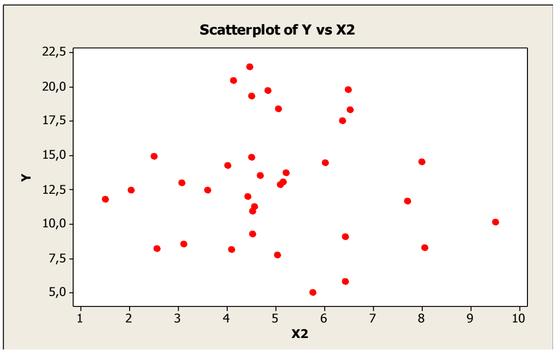
4. HASIL DAN PEMBAHASAN

4.1 *Scatterplot*

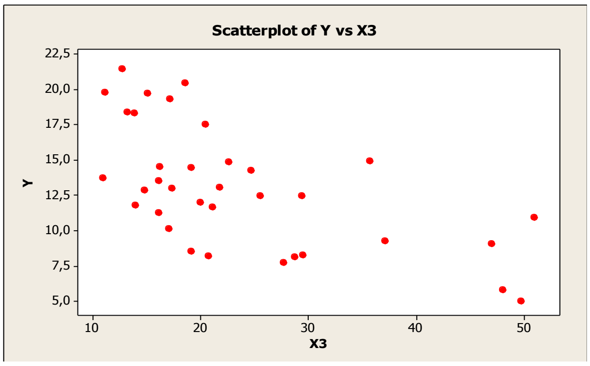
Analisis regresi merupakan salah satu metode statistik yang digunakan untuk mengetahui pola hubungan antara variabel respon dan variabel prediktor. Bentuk pola hubungan antara variabel prediktor dan variabel respon dapat dilihat dari *scatterplot* yang memuat informasi tentang hubungan kedua variabel tersebut.



Gambar 1



Gambar 2



Gambar 3

Ket: *Scatterplot* antara X_1 terhadap Y (Gambar 1), *Scatterplot* antara X_2 terhadap Y (Gambar 2), dan *Scatterplot* antara X_3 terhadap Y (Gambar 3).

Berdasarkan Gambar 1, Gambar 2, dan Gambar 3 plot hubungan antara persentase penduduk miskin dengan laju pertumbuhan ekonomi, tingkat pengangguran terbuka, dan tingkat pendidikan terakhir SMA keatas tidak diketahui bentuk kurva regresinya. Kurva regresi yang diidentifikasi melalui plot-plot yang tersebar menunjukan pola hubungan yang tidak mengikuti suatu pola tertentu.

4.2 Pemilihan Titik Knot Optimal

Model B-spline yang optimal didapatkan dengan menentukan banyak dan letak titik knot dalam beberapa orde. Pendekatan dimulai dari orde 2 sampai orde 4. Banyak titik knot yang akan digunakan untuk setiap orde pada setiap variabel prediktor yaitu satu knot, dua knot dan kombinasi satu dan dua knot sehingga memungkinkan setiap variabel mempunyai banyak knot dan orde yang tidak sama. Letak titik knot juga merupakan hal yang penting untuk mendapatkan model terbaik. Pada penelitian ini, metode yang digunakan untuk menentukan titik knot optimal adalah metode *Generalized Cross Validation* (GCV). Titik knot optimal adalah titik knot dengan GCV minimum, yang nantinya akan menghasilkan estimasi terbaik.

4.3 Model B-spline Terbaik

Nilai GCV optimal untuk masing-masing kombinasi banyak knot pada setiap variabel disajikan kembali pada Tabel 2 berikut:

Tabel 2. Perbandingan Nilai GCV Minimum

Orde			**Titik Knot**			**GCV**
X_1	**X_2**	**X_3**	**X_1**	**X_2**	**X_3**	
2	2	4	4,15212	3,93030	11,41293	11,3088
2	**2**	**2**	**4,51273**	**3,60626**	**11,41293;16,24808**	**9,79353**
2	2	4	4,36849	4,01131;4,09232	11,41293	11,8926
2	2	2	4,32040	4,01131;4,09232	11,41293;16,24808	10,2013
2	2	4	4,39253;4,41657	3,93030	11,41293	11,3088
2	2	2	4,39253;4,41657	3,76828	11,41293;16,24808	9,82981
2	2	4	4,15212;4,17616	4,01131;4,09232	11,41293	11,8926
2	2	2	4,17616;4,20020	3,20121;3,26323	11,4129;16,2481	10,5107

Berdasarkan Tabel 2, estimasi terbaik diperoleh pada saat X_1 berorde 2, X_2 berorde 2, dan X_3 berorde 2 dan banyaknya knot X_1 sebanyak 1 knot yaitu pada titik 4,51273, banyaknya knot X_2 sebanyak 1 knot yaitu pada titik 3,60626, dan banyaknya knot X_3 sebayak 2 knot yaitu pada titik 11,4129 dan 16,2481. Pemilihan model B-spline terbaik dilihat dari nilai GCV minimum dari semua kombinasi orde dan banyak titik knot pada setiap variabel prediktor.

Estimasi model B-spline dengan X_1 berorde 2, X_2 berorde 2, dan X_3 berorde 2 dan banyaknya knot X_1 sebanyak 1 knot, X_2 sebanyak 1 knot, dan X_3 sebanyak 2 knot adalah

$$\hat{y}_i = \hat{\beta}_{11}B_{-1,2}(x_1) + \hat{\beta}_{12}B_{0,2}(x_1) + \hat{\beta}_{13}B_{1,2}(x_1) + \hat{\beta}_{21}B_{-1,2}(x_2) + \hat{\beta}_{22}B_{0,2}(x_2)$$
$$+\hat{\beta}_{23}B_{1,2}(x_2) + \hat{\beta}_{31}B_{-1,2}(x_3) + \hat{\beta}_{32}B_{0,2}(x_3) + \hat{\beta}_{33}B_{1,2}(x_3) + \hat{\beta}_{34}B_{2,2}(x_3)$$

Berdasarkan *running* program menggunakan R 3.0.2, diperoleh estimasi parameter $\hat{\beta}$ untuk model B-spline terbaik yang dapat dilihat pada Tabel 3 sebagai berikut:

Tabel 3. Estimasi Parameter Model B-spline Terbaik

Variabel	Parameter	Estimasi Parameter
X_1	β_{11}	1,14495
	β_{12}	1,57147
	β_{13}	9,622754
X_2	β_{21}	1,763506
	β_{22}	7,489224
	β_{23}	3,086443
X_3	β_{31}	1,373524
	β_{32}	12,416796
	β_{33}	1,900542
	β_{34}	-3,351689

Berdasarkan Tabel 3. Estimasi model B-spline terbaik dapat dituliskan sebagai berikut:

$$\begin{aligned}\hat{y}_i &= 1{,}14495\, B_{-1,2}(x_1) + 1{,}57147\, B_{0,2}(x_1) + 9{,}622754\, B_{1,2}(x_1) + 1{,}763506\, B_{-1,2}(x_2)\\ &+7{,}489224\, B_{0,2}(x_2) + 3{,}086443\, B_{1,2}(x_2) + 1{,}373524\, B_{-1,2}(x_3)\\ &+12{,}416796\, B_{0,2}(x_3) + 1{,}900542\, B_{1,2}(x_3) - 3{,}351689\, B_{2,2}(x_3)\end{aligned}$$

dengan:

$$B_{-1,2}(x_1) = \begin{cases} \dfrac{4{,}51273 - x_1}{0{,}43273}, & 4{,}08 \le x_1 \le 4{,}51273 \\ 0, & untuk\ lainya \end{cases}$$

$$B_{0,2}(x_1) = \begin{cases} \dfrac{x_1 - 4{,}08}{0{,}43273}, & 4{,}08 \le x_1 \le 4{,}51273 \\ \dfrac{6{,}46 - x_1}{1{,}94727}, & 4{,}51273 \le x_1 \le 6{,}46 \\ 0, & untuk\ lainya \end{cases}$$

$$B_{1,2}(x_1) = \begin{cases} \dfrac{x_1 - 4{,}51273}{1{,}942727}, & 4{,}51273 \le x_1 \le 6{,}46 \\ 0, & untuk\ lainya \end{cases}$$

$$B_{-1,2}(x_2) = \begin{cases} \dfrac{3{,}60626 - x_2}{2{,}10628}, & 1{,}5 \le x_2 \le 3{,}60626 \\ 0, & untuk\ lainya \end{cases}$$

$$B_{0,2}(x_2) = \begin{cases} \dfrac{x_2 - 1{,}5}{2{,}10626}, & 1{,}5 \le x_2 \le 3{,}60626 \\ \dfrac{9{,}52 - x_2}{5{,}91374}, & 3{,}60626 \le x_2 \le 9{,}52 \\ 0, & untuk\ lainya \end{cases}$$

$$B_{1,2}(x_2) = \begin{cases} \dfrac{x_1 - 3{,}60626}{5{,}91374}, & 3{,}60626 \le x_2 \le 9{,}52 \\ 0, & untuk\ lainya \end{cases}$$

$$B_{-1,2}(x_3) = \begin{cases} \dfrac{11{,}4129 - x_3}{0{,}4029}, & 11{,}01 \le x_3 \le 11{,}4129 \\ 0, & untuk\ lainya \end{cases}$$

$$B_{0,2}(x_3) = \begin{cases} \dfrac{x_3 - 11{,}01}{0{,}4029} & , \ 11{,}01 \leq x_3 \leq 11{,}4129 \\ \dfrac{16{,}2481 - x_3}{4{,}8352} & , \ 11{,}4129 \leq x_3 \leq 16{,}2481 \\ 0 & , \ untuk\ lainya \end{cases}$$

$$B_{1,2}(x_3) = \begin{cases} \dfrac{x_3 - 11{,}4129}{4{,}8352} & , \ 11{,}4129 \leq x_3 \leq 16{,}2481 \\ \dfrac{50{,}9 - x_3}{34{,}6519} & , \ 16{,}2481 \leq x_3 \leq 50{,}9 \\ 0 & , \ untuk\ lainya \end{cases}$$

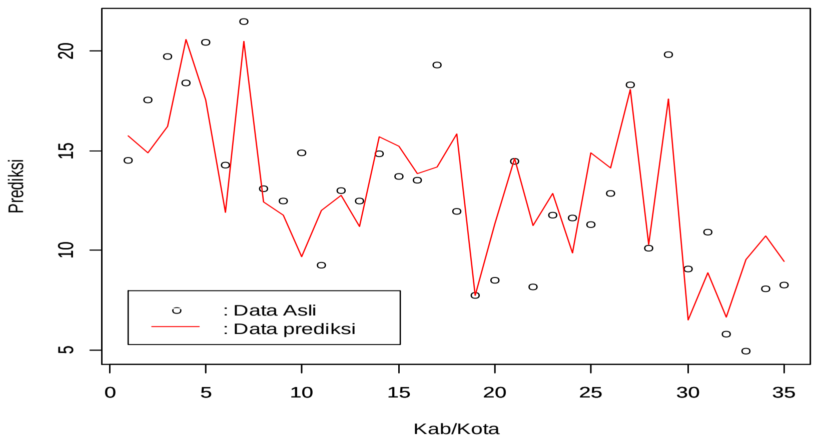
dan

$$B_{2,2}(x_3) = \begin{cases} \dfrac{x_3 - 16{,}2481}{34{,}6519} & , \ 16{,}2481 \leq x_3 \leq 50{,}9 \\ 0 & , \ untuk\ lainya \end{cases}$$

Berdasarkan *running* pada program R 3.0.2, diperolah nilai $R^2 = 0{,}6778968$. Nilai R^2 tersebut menunjukkan bahwa pengaruh variabel laju pertumbuhan ekonomi, variabel tingkat pengangguran terbuka dan variabel tingkat pendidikan terakhir SMA ke atas terhadap persentase penduduk miskin sebesar 67,79%.

4.4 Komparasi Data Asli dengan Hasil Prediksi

Setelah didapatkan model B-spline terbaik yaitu pada saat variabel X_1 berorde 2, variabel X_2 berorde 2, dan variabel X_3 berorde 2 dan banyaknya knot X_1 sebanyak 1 knot, X_2 sebanyak 1 knot, dan X_3 sebayak 2 knot, maka dapat diperoleh nilai prediksinya. Kurva estimasi data asli dengan data prediksi dapat dilihat pada kurva berikut:



Gambar 4. Kurva Estimasi Data Asli dan Data Prediksi

Berdasarkan Gambar 4, dapat dilihat bahwa data prediksi yang dihasilkan cenderung mengikuti pergerakan data aslinya, meskipun kurva tersebut menunjukan bahwa nilai prediksinya tidak sama persis dengan nilai aktualnya. Hal tersebut berarti bahwa model regresi B-spline yang diperoleh mempunyai kemampuan menyesuaikan diri lebih efektif dalam mengatasi pola data yang naik atau turun yang tajam dengan bantuan titik-titik knot.

5 KESIMPULAN

Berdasarkan analisis yang telah dilakukan, maka dapat diambil kesimpulan sebagai berikut:

1. Model B-spline terbaik untuk estimasi data persentase penduduk miskin di Jawa Tengah adalah model B-spline pada saat X_1 berorde 2, X_2 berorde 2, dan X_3 berorde 2 dan banyaknya knot X_1 sebanyak 1 knot yaitu pada titik 4,51273, banyaknya knot X_2 sebanyak 1 knot yaitu pada titik 3,60626, dan banyaknya knot X_3 sebanyak 2 knot yaitu pada titik 11,4129 dan 16,2481 dengan nilai GCV sebesar 9,79353.
2. Persamaan model B-spline terbaiknya sebagai berikut:

$$\begin{aligned}\hat{y}_i = {} & 1{,}144951\, B_{-1,2}(x_1) + 1{,}571462\, B_{0,2}(x_1) + 9{,}622755\, B_{1,2}(x_1) \\ & +1{,}763508\, B_{-1,2}(x_2) + 7{,}489221\, B_{0,2}(x_2) + 3{,}08644\, B_{1,2}(x_2) \\ & +1{,}373525\, B_{-1,2}(x_3) + 12{,}416791\, B_{0,2}(x_3) + 1{,}900542\, B_{1,2}(x_3) \\ & -3{,}35169\, B_{2,2}(x_3)\end{aligned}$$

DAFTAR PUSTAKA

BPS. 2016. *Data dan Informasi Kemiskinan Kabupaten/Kota 2015*. Badan Pusat Statistik, Jakarta.

BPS. 2016. *Keadaan Angkatan Kerja di Jawa Tengah 2015*. BPS Provinsi Jawa Tengah, Semarang.

BPS. 2016. *Produk Domestik Regional Bruto Kabupaten/Kota di Indonesia 2011- 2015*. Badan Pusat Statistik, Jakarta.

BPS. 2016. *Statistik Sosial dan Kependudukan Jawa Tengah 2014*. BPS Provinsi Jawa Tengah, Semarang.

Budiantara, I.N., Suryadi, F., Otok, B.W., dan Guritno, S., 2006. *Permodelan B-Spline dan MARS Pada Nilai Ujian Masuk Terhadap IPK Mahasiswa Jurusan Disain Komunikasi Visual UK. PETRA Surabaya*. Jurnal Teknik Industri. 8(1): 1-13.

Eubank, R. L. 1999. *Nonparametric Regression and Spline Smoothing*, 2nd ed. New York: Marcel Dekker.

Gujarati, D. 2004. *Basic Econometrics*, *Fourth Edition*. New York: The McGraw−Hill Companies.

Merdekawati, I. P., dan Budiantara, I. N. 2013. *Pemodelan Regresi spline Truncated Multivariabel pada Faktor-faktor yang mempengaruhi Kemiskinan di Kabupaten/Kota Provinsi Jawa Tengah*. Jurnal Sains dan Seni Pomits Vol. 2, No. 1, 2337-3520.

Prastyo, A. A. 2010. *Analisis Faktor-faktor yang Mempengaruhi Tingkat Kemiskinan (Studi Kasus 35 Kabupaten/Kota di Jawa Tengah Tahun 2003-2007)*. Skripsi Sarjana, Fakultas Ekonomika dan Bisnis, Universitas Diponegoro, Semarang.

Winarendra, A. 2014. *Analisis Tingkat Kemiskinan dan Faktor-faktor yang Mempengaruhi (Kasus: 35 Kabupaten/Kota di Jawa Tengah 2008-2012).* Skripsi Sarjana, Fakultas Ekonomika dan Bisnis, Universitas Diponegoro, Semarang.