

REGRESI NONPARAMETRIK *SPLINE* PADA DATA LAJU PERTUMBUHAN EKONOMI DI KALIMANTAN

Spline Nonparametric Regression on Economic Growth Rate Data in Kalimantan

Tirta Purnaraga¹, Sifriyani^{2*}, Surya Prangga³

¹Program Studi Statistika ²Fakultas Matematika dan Ilmu Pengetahuan Alam ³Universitas Mulawarman
Jln. Barong Tongkok No.4 Kampus Gn. Kelua, Samarinda, 75123, Indonesia

e-mail: ¹tirtapurnaraga@gmail.com ; ^{2*}sifri.stat.unmul@gmail.com ; ³suryapringga@gmail.com
Corresponding author*

Abstrak

Laju Pertumbuhan Ekonomi (LPE) merupakan indikator penting untuk mengukur keberhasilan pembangunan suatu perekonomian. Kesejahteraan dan kemajuan suatu perekonomian ditentukan oleh besarnya pertumbuhan yang ditunjukkan oleh perubahan kuantitas produksi barang dan jasa secara nasional. Pertumbuhan ekonomi yang tinggi merupakan sasaran yang diharapkan dapat tercapai di sebuah negara berkembang. Kalimantan merupakan salah satu regional di Indonesia yang kontribusinya terhadap pertumbuhan ekonomi nasional adalah terbesar ketiga setelah Jawa dan Sumatera pada tahun 2017. Banyak faktor yang mempengaruhi LPE di Kalimantan sehingga perlu dilakukan pemodelan untuk mengetahui faktor-faktor yang secara signifikan mempengaruhi LPE. Penelitian ini menggunakan 6 faktor yang diduga mempengaruhi LPE yaitu tingkat partisipasi angkatan kerja, jumlah industri besar dan sedang, rata-rata lama sekolah, anggaran pendapatan dan belanja daerah, dana alokasi umum dan produktivitas padi. Data tersebut merupakan data tahun 2017 yang diperoleh dari Badan Pusat Statistik (BPS) di 5 Provinsi di Kalimantan. Metode yang digunakan untuk memodelkan LPE adalah regresi nonparametrik *spline* dan diperoleh titik knot optimal 3 titik knot berdasarkan nilai *Generalized Cross Validation* (GCV) terkecil 1,208. Berdasarkan hasil penelitian diperoleh model terbaik dengan nilai R^2 sebesar 82,15 persen dan *Mean Square Error* (MSE) sebesar 0,805. Hasil penelitian memberikan informasi bahwa faktor-faktor yang mempengaruhi LPE adalah tingkat partisipasi angkatan kerja, jumlah industri besar dan sedang, rata-rata lama sekolah, anggaran pendapatan dan belanja daerah, dana alokasi umum dan produktivitas padi.

Kata Kunci: Regresi nonparametrik, *spline*, LPE, titik knot, GCV

Abstract

Economic Growth Rate (EGR) is an important indicator for measuring the success of an economy's development. The welfare and progress of an economy is determined by the amount of growth shown by changes in the quantity of goods and services produced nationally. High economic growth is a goal that is expected to be achieved in a developing country. Many factors affect EGR in Kalimantan, so it is necessary to do modeling to find out the factors that significantly affect EGR. This study uses 6 factors that are suspected to influence EGR, namely the labor force participation rate, the number of large and medium industries, the average length of schooling, regional income and expenditure budgets, general allocation funds and rice productivity. The data is 2017 data obtained from the Central Bureau of Statistics in 5 provinces in Kalimantan. The method used to model the LPE is spline nonparametric regression and the optimal knot point is 3 knot points based on the smallest Generalized Cross Validation (GCV) value of 1.208. The research results, the best model is obtained with a R^2 value of 82.15 percent and a Mean Square Error (MSE) of 0.805. The results of the study provide information that the factors that influence the LPE are the level of labor force participation, the number of large and medium industries, the average length of schooling, regional income and expenditure budgets, general allocation funds and rice productivity.

Keywords: Nonparametric regression, *spline*, LPE, knot point, GCV

Submitted: 26th May 2020

Accepted: 12th July 2020

1. PENDAHULUAN

Analisis regresi merupakan salah satu analisis dalam statistika yang digunakan untuk menyelidiki pola hubungan fungsional antara variabel respon dan variabel prediktor (mencari bentuk estimasi kurva regresi). Regresi nonparametrik bentuk kurva regresinya tidak diketahui, data diharapkan mencari sendiri bentuk estimasinya sehingga memiliki fleksibilitas yang tinggi. Data yang memiliki pola yang berubah-ubah pada sub-sub interval tertentu sebaiknya dimodelkan menggunakan metode *spline*. Hal yang perlu diperhatikan dalam membentuk model pada regresi *spline* yaitu menentukan orde model, banyaknya titik knot dan lokasi titik knot tersebut [1]. Orde model dapat diketahui dari bentuk pola antar variabel. Hal yang tidak kalah pentingnya adalah menentukan banyak titik knot dan lokasi dari tiap-tiap knot karena akan berpengaruh pada bentuk kurva regresi. Ada beberapa metode yang dapat digunakan untuk memilih titik knot optimal dalam estimator *spline*, antara lain metode *Cross Validation* (CV), *Unbiased Risk* (UBR), *Generalized Maximum Likelihood* (GML) dan *Generalized Cross Validation* (GCV) [14].

Laju pertumbuhan ekonomi adalah ukuran pertumbuhan ekonomi dari satu periode ke periode lain menggunakan persentase. Ukuran ini tidak menyesuaikan inflasi, namun dinyatakan dalam istilah nominal. Dalam penelitian ini, studi kasusnya berkaitan dengan laju pertumbuhan ekonomi di Indonesia. Perekonomian Indonesia tahun 2017 tumbuh 5,07 persen, pencapaian ini lebih tinggi dibanding pertumbuhan ekonomi tahun 2016 yaitu sebesar 5,03 persen. Dari sisi produksi, pertumbuhan tertinggi dicapai oleh Lapangan Usaha Informasi dan Komunikasi sebesar 9,81 persen. Dari sisi pengeluaran pertumbuhan tertinggi dicapai oleh Komponen Ekspor Barang dan Jasa sebesar 9,09 persen. Apabila kita bandingkan dengan laju pertumbuhan ekonomi regional semisal Pulau Kalimantan yang terdiri dari 5 provinsi yaitu Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kalimantan Utara, dan Kalimantan Timur [2].

Wilayah Kalimantan Barat mengalami pertumbuhan sebesar 5,80 persen pada tahun 2017, didorong oleh hampir semua lapangan usaha dan sisi pengeluaran. Pada wilayah Provinsi Kalimantan Selatan, perekonomian tumbuh 5,29 persen selama tahun 2017, Pertumbuhan ekonomi tersebut hampir didukung oleh semua sektor [3]. Pada wilayah Kalimantan Tengah pertumbuhan Ekonomi pada tahun 2017 tumbuh 6,74 persen. Struktur perekonomian Kalimantan Tengah tahun 2017 didominasi oleh tiga kategori utama yakni pertanian, kehutanan dan perikanan [4]. Kalimantan Utara yang pertumbuhan ekonominya tumbuh hampir 2 kali lipat dimana pada tahun 2017 tumbuh sebesar 6,59 persen, didominasi oleh tiga lapangan usaha utama yaitu lapangan usaha pertambangan dan penggalian dengan peranan sebesar 27,38 persen, lapangan usaha pertanian, kehutanan, dan perikanan dengan peranan sebesar 16,32 persen, dan lapangan usaha konstruksi dengan peranan sebesar 12,45 persen terhadap pembentukan Pendapatan Domestik Regional Bruto (PDRB) Provinsi Kalimantan Utara [5]. Kalimantan Timur ekonominya pada tahun 2017 tumbuh sebesar 3,13 persen meski terkesan kecil namun sebenarnya ini adalah pencapaian besar karena pada tahun 2016 pertumbuhan ekonomi minus 0,38 persen. Ekspor luar negeri masih mendominasi perekonomian Kalimantan Timur. Meski secara regional tiap provinsi sebagian besar tumbuh melebihi pencapaian ekonomi nasional, tetapi kontribusi wilayah Kalimantan terhadap pertumbuhan ekonomi nasional pada 2017 hanya 8,2 persen [6].

Beberapa penelitian sebelumnya yang menggunakan regresi nonparametrik *spline*, yaitu “Analisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka di Kalimantan Menggunakan Regresi Nonparametrik *Spline Truncated*” [7], “*Nonparametric Geographically Weighted Regression* dengan Pendekatan *Spline Truncated* pada Data Produktivitas Tanaman Padi” [8], “Evaluasi Faktor yang Mempengaruhi Laju Pertumbuhan Ekonomi di Kalimantan Menggunakan Regresi Nonparametrik *Spline* dengan Pembobot Geografis” [9], dan “Pemodelan Regresi Nonparametrik *Spline Truncated* dan Aplikasinya Pada Indeks Pembangunan Manusia di Pulau Kalimantan” [10].

Berdasarkan uraian dari latar belakang diatas maka penulis tertarik melakukan suatu penelitian untuk mengetahui faktor-faktor yang mempengaruhi laju pertumbuhan ekonomi. Hal ini bertujuan untuk menyelidiki faktor-faktor yang diduga mempengaruhi laju pertumbuhan ekonomi di Kalimantan.

2. METODE PENELITIAN

2.1 Regresi Nonparametrik

Pendekatan regresi nonparametrik menyesuaikan dengan data yang memiliki pola yang tidak diketahui bentuk kurva regresinya atau tidak terdapat informasi masa lalu yang lengkap tentang bentuk pola data (Eubank, 1998). Regresi nonparametrik mulai dikenal sekitar abad 19, tepatnya pada tahun 1857 (Hardle, 1990). Secara visual, jenis regresi ini memiliki bentuk pola yang tidak dapat diidentifikasi. Pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi, karena data diharapkan mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi faktor subyektifitas peneliti [11].

Regresi nonparametrik memiliki fleksibilitas yang tinggi karena bentuk estimasi kurva regresinya dapat menyesuaikan dengan tanpa dipengaruhi oleh faktor subyektifitas peneliti [12]. Selanjutnya $f(x_i)$ merupakan kurva regresi yang diasumsikan *smooth* dalam arti merupakan anggota fungsi tertentu. Bentuk kurva regresi $f(x_i)$ diasumsikan tidak diketahui dan termuat di dalam ruang sobolev $W_2^m[a, b]$ khususnya $f(x_i) \in W_2^m[a, b]$ dengan $W_2^m[a, b] = \{f: f, f^{(1)}, f^{(2)}, \dots, f^{(m-1)} \text{ absolute continues pada } [a, b], \int_a^b [f^{(m)}(x)]^2 dx < \infty\}$ untuk suatu m bilangan bulat positif [13].

2.2 Regresi Nonparametrik dengan Pendekatan Spline

Spline merupakan model polinomial yang tersegmen. Polinomial tersegmen memegang peranan penting dalam teori dan aplikasi statistika. Sifat tersegmen ini memberikan sifat fleksibilitas yang lebih baik dari polinomial biasa dan sifat ini memungkinkan model regresi *spline* untuk menyesuaikan diri lebih efektif terhadap karakteristik lokal dari data. *Spline* sangat tergantung pada titik knot. Titik knot merupakan titik perpaduan bersama dimana terjadi pola perubahan perilaku dari suatu fungsi pada selang yang berbeda [14].

Secara umum fungsi f dalam ruang *spline* berorde m dengan titik knot k_1, k_2, \dots, k_h adalah sembarang fungsi yang dapat dinyatakan sebagai berikut.

$$f(x_i) = \sum_{k=0}^m \beta_k x_i^k + \sum_{h=1}^r \beta_{m+h} (x_i - k_h)_+^m$$

dengan:

$$(x_i - k_h)_+^m = \begin{cases} (x_i - k_h)^m, & x_i \geq k_h \\ 0, & x_i < k_h \end{cases} \tag{1}$$

β_k merupakan parameter-parameter model dengan m merupakan orde *spline*. Model pada Persamaan (1) untuk setiap pengamatan $i = 1, 2, \dots, n$ dapat dituliskan pada Persamaan (2).

$$\begin{aligned} y_i &= \sum_{k=0}^m \beta_k x_i^k + \sum_{h=1}^r \beta_{m+h} (x_i - k_h)_+^m + \varepsilon_i \\ &= \beta_0 + \beta_1 x_i + \dots + \beta_m x_i^m + \beta_{m+1} (x_i - k_1)_+^m + \beta_{m+2} (x_i - k_2)_+^m + \dots + \\ &\quad \beta_{m+r} (x_i - k_r)_+^m + \varepsilon_i \end{aligned} \tag{2}$$

Secara matematis hubungan antara variabel respon y_i dan variabel prediktor $x_{1i}, x_{2i}, \dots, x_{Li}$ untuk model regresi nonparametrik multivariabel *spline* adalah

$$\begin{aligned} y_i &= \sum_{p=1}^L \left(\sum_{k=0}^m \beta_{pk} x_{pi}^k + \sum_{h=1}^r \beta_{p,m+h} (x_{pi} - k_{ph})_+^m \right) \\ &= \sum_{p=1}^L \sum_{k=0}^m \beta_{pk} x_{pi}^k + \sum_{h=1}^r \beta_{p,m+h} (x_{pi} - k_{ph})_+^m \\ &= \beta_{10} + \beta_{11} x_{1i} + \beta_{12} x_{1i}^2 + \beta_{13} x_{1i}^3 + \dots + \beta_{1m} x_{1i}^m + \beta_{1,m+1} (x_1 - k_{11})_+^m + \beta_{1,m+2} (x_1 - k_{12})_+^m + \\ &\quad \dots + \beta_{1,m+h} (x_1 - k_{1h})_+^m + \beta_{20} + \beta_{21} x_{2i} + \beta_{22} x_{2i}^2 + \beta_{23} x_{2i}^3 + \dots + \beta_{2m} x_{2i}^m + \beta_{2,m+1} (x_2 - \\ &\quad k_{21})_+^m + \beta_{2,m+2} (x_2 - k_{22})_+^m + \dots + \beta_{2,m+h} (x_2 - k_{2h})_+^m + \dots + \beta_{L,m+r} (x_L - k_{Lr})_+^m + \varepsilon_i \end{aligned} \tag{3}$$

Karena $\beta_{10}, \dots, \beta_{L0}$ adalah suatu konstanta sehingga Persamaan (3) dapat disederhanakan sebagai berikut:

$$y_i = \beta_0 + \sum_{p=1}^L \sum_{k=1}^m \beta_{pk} x_{pi}^k + \sum_{p=1}^L \sum_{h=1}^r \beta_{p,m+h} (x_{pi} - k_{ph})_+^m + \varepsilon_i \tag{4}$$

Persamaan (4) dapat dituliskan dalam bentuk matriks sebagai berikut:

$$\mathbf{y} = \mathbf{X}(\mathbf{k})\boldsymbol{\beta}$$

Dimana:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{1m} \\ \beta_{21} \\ \beta_{22} \\ \vdots \\ \beta_{2m} \\ \vdots \\ \beta_{pm} \\ \beta_{pm+1} \\ \vdots \\ \beta_{pm+r} \end{bmatrix}$$

$$\mathbf{X}(\mathbf{k}) = \begin{bmatrix} 1 & x_{11} & x_{11}^2 & \dots & x_{11}^m & x_{21} & x_{21}^2 & \dots & x_{21}^m & \dots & x_{L1}^m & (x_{11} - k_{11})_+^m & \dots & (x_{p1} - k_{pr})_+^m \\ 1 & x_{12} & x_{12}^2 & \dots & x_{12}^m & x_{22} & x_{22}^2 & \dots & x_{22}^m & \dots & x_{L2}^m & (x_{12} - k_{11})_+^m & \dots & (x_{p2} - k_{pr})_+^m \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{1n}^2 & \dots & x_{1n}^m & x_{2n} & x_{2n}^2 & \dots & x_{2n}^m & \dots & x_{Ln}^m & (x_{1n} - k_{11})_+^m & \dots & (x_{pn} - k_{pr})_+^m \end{bmatrix}$$

Persamaan (4) mempunyai vektor \mathbf{y} dengan ukuran $n \times 1$, matriks $\mathbf{X}(\mathbf{k})$ dengan ukuran $n \times (\mathbf{1} + \mathbf{L} + \mathbf{r})$ dan vektor $\boldsymbol{\beta}$ dengan ukuran $(\mathbf{1} + \mathbf{L} + \mathbf{r}) \times 1$.

2.3 Estimasi Parameter Regresi Spline

Estimator *spline* dalam regresi nonparametrik mampu menangani data/fungsi yang mulus (*smooth*) [15]. Kemampuan ini ditunjukkan oleh fungsi *truncated* (potongan-potongan) yang melekat pada estimator tersebut. Apabila pada persamaan (3) diasumsikan error $\varepsilon_i \sim \text{IIDN}(0, \sigma^2)$, maka kepadatan bersamanya adalah sebagai berikut.

$$f(y_1, y_2, \dots, y_n) = \prod_{i=1}^n \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2\sigma^2} (y_i - \mu)^2 \right] \right]$$

$$= \prod_{i=1}^n \left[(2\pi)^{-1/2} (\sigma^2)^{-1/2} \exp \left(-\frac{1}{2\sigma^2} (y_i - \mu)^2 \right) \right]$$

Akibatnya diperoleh fungsi *likelihood* sebagai berikut.

$$L(\beta, \sigma^2 | y) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(x_i))^2 \right) \quad (5)$$

Estimasi titik untuk fungsi f diperoleh dengan menyelesaikan optimasi *likelihood* sebagai berikut

$$\text{Max}_{\beta} L(\beta, \sigma^2 | y) = \text{Max}_{\beta} \left\{ (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(x_i))^2 \right) \right\} \quad (6)$$

Apabila diambil Transformasi Logaritma dan mengingat Persamaan (1) maka diperoleh sebagai berikut.

$$\text{Ln } L(\beta, \sigma^2 | y) = \text{Ln} \left\{ (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(x_i))^2 \right) \right\}$$

$$\begin{aligned}
 &= -\frac{n}{2} \text{Ln}(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(x_i))^2 \\
 &= -\frac{n}{2} \text{Ln}(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n \left(y_i - \left(\sum_{k=0}^m \beta_k x_i^k + \sum_{h=1}^r \beta_{m+h} (x_i - k_h)_+^m \right) \right)
 \end{aligned} \tag{7}$$

Untuk regresi nonparametrik *spline* multivariabel, fungsi *likelihood*nya adalah :

$$\text{Ln } L(\beta, \sigma^2 | y) = -\frac{n}{2} \text{Ln}(2\pi\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}(\mathbf{k})\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}(\mathbf{k})\boldsymbol{\beta}) \tag{8}$$

Jika Persamaan (8) diturunkan terhadap β , kemudian hasilnya disamakan dengan nol, didapatkan hasil sebagai berikut.

$$\frac{\partial \text{Ln } L(\beta, \sigma^2 | y)}{\partial \beta} = \frac{\partial}{\partial \beta} \left[-\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}(\mathbf{k})\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}(\mathbf{k})\boldsymbol{\beta}) \right] = 0 \tag{9}$$

Dari penjabaran persamaan (9), maka diperoleh estimasi *Likelihood* untuk model regresi $\beta(\mathbf{k})$ yang diberikan oleh Persamaan (10).

$$\hat{\beta}(\mathbf{k}) = \mathbf{X}(\mathbf{k})[\mathbf{X}(\mathbf{k})^T \mathbf{X}(\mathbf{k})]^{-1} \mathbf{y} \tag{10}$$

Selanjutnya, dalam menentukan titik knot optimal diperlukan metode untuk mendapatkan model *spline* terbaik, salah satunya adalah metode GCV.

2.4 Pemilihan Titik Knot Optimal

Titik knot itu sendiri merupakan titik perpaduan bersama dimana terdapat perubahan perilaku pola pada interval yang berlainan. Wahba (1990) memberikan suatu metode yang baik untuk memilih titik knot optimal yaitu metode GCV. Secara teoritis bahwa metode GCV mempunyai sifat optimal asimtotik. Disamping metode GCV, juga bisa menggunakan metode CV. Metode CV dilakukan dengan menghapus setiap titik data secara bergantian kemudian dijumlahkan dan dikuadratkan dari perbedaan dari setiap titik data yang dihilangkan dari data yang cocok ke semua titik data lainnya. Rumus dari Metode GCV didefinisikan pada Persamaan (11).

$$\text{GCV}(\mathbf{k}) = \frac{\text{MSE}(\mathbf{k})}{(n^{-1} \text{trace}[\mathbf{I} - \mathbf{B}(\mathbf{k})])^2} \tag{11}$$

Pada persamaan (11) \mathbf{I} adalah matriks identitas, n adalah jumlah pengamatan dan $\mathbf{B}(\mathbf{k})$ adalah matriks $\mathbf{X}(\mathbf{k})[\mathbf{X}(\mathbf{k})^T \mathbf{X}(\mathbf{k})]^{-1} \mathbf{X}(\mathbf{k})^T$ sedangkan nilai $\text{MSE}(\mathbf{k})$ diberikan pada Persamaan (12).

$$\text{MSE}(\mathbf{k}) = n^{-1} \sum_{i=1}^n (y_i - f(x_i))^2 \tag{12}$$

2.5 Uji Simultan

Uji simultan merupakan pengujian parameter model regresi secara bersamaan. Uji ini bertujuan untuk mengetahui parameter model regresi telah signifikan atau tidak. Hipotesis yang digunakan adalah sebagai berikut:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_m = 0$$

$$H_1: \text{minimal terdapat satu } \beta_j \neq 0, j = 1, 2, \dots, m.$$

Nilai m adalah jumlah parameter yang terdapat dalam model regresi dan n merupakan jumlah observasi.

Tabel 1. Analisis Variansi

Sumber Variasi	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Batasan Kuadrat (MS)	F_{hitung}
Regresi	$(m + J - 1)$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SS_{Regresi}}{df_{Regresi}}$	$\frac{MS_{regresi}}{MS_{residual}}$
Residual	$n - (m + J)$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SS_{Residual}}{df_{Residual}}$	
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$		

Daerah penolakan H_0 jika nilai $F_{hitung} > F_{\alpha((m-1),(n-m))}$. Apabila pada kesimpulan dihasilkan penolakan H_0 , maka dapat dinyatakan bahwa terdapat minimal satu parameter $\beta_j \neq 0$. Hal tersebut mengindikasikan bahwa terdapat minimal satu variabel prediktor yang berpengaruh terhadap variabel respon.

Salah satu kriteria untuk mengukur kebaikan model adalah dengan menggunakan R^2 . Dimana jika nilai $R^2 > 70\%$ maka dapat dikatakan model telah dianggap baik (Walpole, 1995). Berikut adalah persamaan untuk mendapatkan nilai R^2 .

$$R^2 = \frac{SS_{regresi}}{SS_{total}} \quad (13)$$

2.6 Uji Individu

Jika pada uji simultan disimpulkan bahwa minimal terdapat satu parameter β_j signifikan, maka perlu diketahui secara individu parameter mana yang signifikan dan parameter mana yang tidak signifikan. Hal ini dapat dilakukan melalui uji individu. Rumusan hipotesisnya sebagai berikut.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, m$$

Statistik uji sebagai berikut.

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} = \frac{\hat{\beta}_j}{\sqrt{\sigma^2 C_{jj}}} \quad (14)$$

Berdasarkan Persamaan (12) C_{jj} adalah elemen diagonal ke- j dari invers matriks $(X^T X)^{-1}$ dan $SE(\hat{\beta}_j) = \sqrt{\sigma^2 C_{jj}}$. Daerah penolakan H_0 adalah $|t_{hitung}| > t_{(\alpha/2, (n-m))}$.

2.7 Sumber Data

Sumber data pada penelitian ini adalah data sekunder yang diambil dari website resmi Badan Pusat Statistik (BPS) di tiap provinsi regional Kalimantan pada tahun 2017. Penelitian ini menggunakan 1 variabel respon dan 6 variabel prediktor. Variabel responnya adalah laju pertumbuhan ekonomi dan variabel prediktornya adalah tingkat partisipasi angkatan kerja, rata-rata lama sekolah, jumlah perusahaan industri besar dan sedang, produktivitas padi, anggaran pendapatan dan belanja daerah dan dana alokasi umum.

Tabel 2. Sumber Data

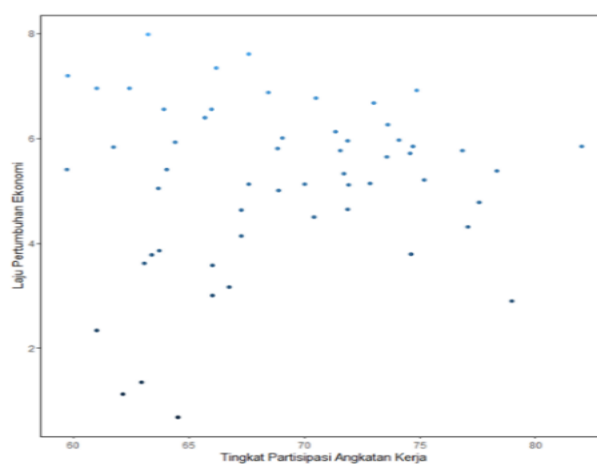
Variabel		Keterangan	Satuan
Respon (y)	(y)	Laju pertumbuhan ekonomi di 56 kab/kota di Kalimantan	persen
	(x_1)	Tingkat partisipasi angkatan kerja di 56 kab/kota di Kalimantan	persen
Prediktor (x)	(x_2)	Jumlah industri besar dan sedang di 56 kab/kota di Kalimantan	unit
	(x_3)	Rata-rata lama sekolah di 56 kab/kota di Kalimantan	tahun

(x_4)	Anggaran Pendapatan dan Belanja Daerah di 56 Kab/Kota di Kalimantan	triliun rupiah
(x_5)	Dana Alokasi Umum di 56 Kab/Kota di Kalimantan	triliun rupiah
(x_6)	Produktivitas Padi di 56 Kab/Kota di Kalimantan	kwintal/hektar

3. HASIL DAN PEMBAHASAN

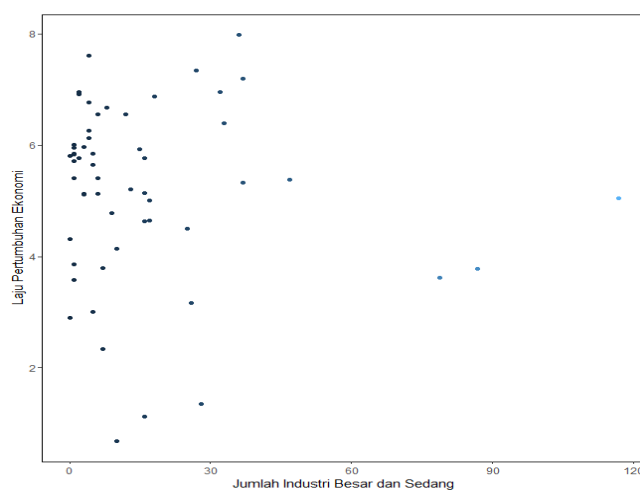
Langkah pertama dibentuk *scatterplot* kemudian dilanjutkan dengan mencari titik knot optimal kemudian dilakukan pengujian signifikansi model secara simultan, bersamaan dengan itu juga dilakukan pengujian signifikansi model secara individu hingga kemudian sampai pada interpretasi model.

3.1 Scatterplot



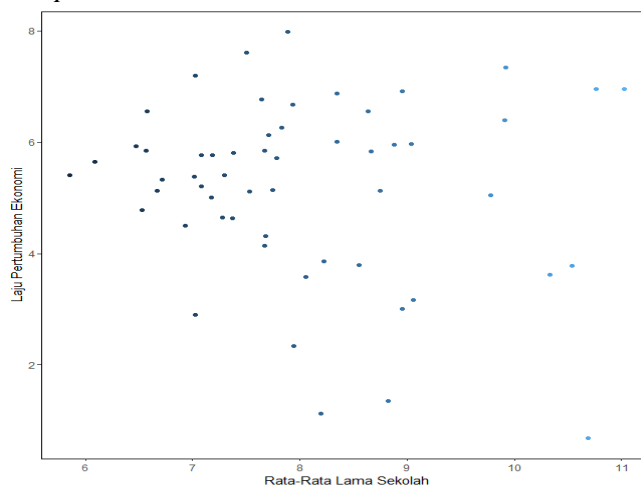
Gambar 1. Scatterplot antara LPE dengan Tingkat Partisipasi Angkatan Kerja

LPE diduga dipengaruhi oleh Tingkat Partisipasi Angkatan Kerja (TPAK), dimana TPAK menunjukkan proporsi penduduk usia kerja yang terlibat secara aktif dalam pasar tenaga kerja baik yang bekerja maupun yang sedang mencari pekerjaan. Semakin tinggi TPAK di suatu kabupaten/kota maka peluang untuk memperoleh pekerjaan dan mengurangi tingkat pengangguran juga semakin tinggi. Pola hubungan antara LPE dengan tingkat partisipasi angkatan kerja (x_1) terlihat tidak membentuk suatu pola tertentu seperti pada Gambar 1. Oleh sebab itu, metode yang digunakan adalah regresi nonparametrik *spline*.



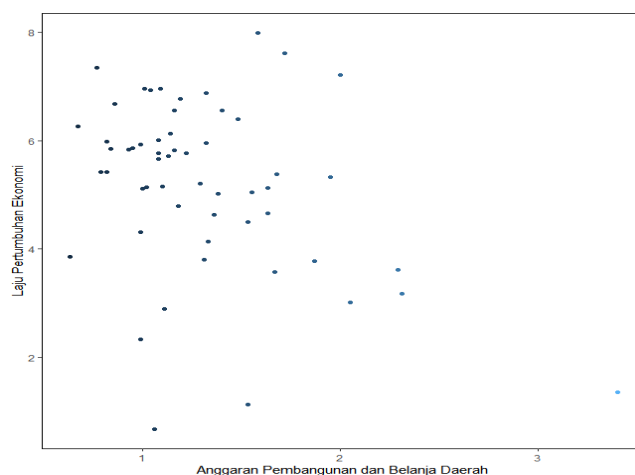
Gambar 2. Scatterplot antara LPE dengan Jumlah Industri Besar Sedang

LPE diduga dipengaruhi oleh jumlah industri besar sedang, dimana jumlah industri besar sedang menunjukkan banyaknya industri besar sedang di suatu daerah. Semakin tinggi jumlah industri besar sedang di suatu Kabupaten/Kota maka peluang untuk memperoleh pekerjaan dan mengurangi tingkat pengangguran juga semakin tinggi. Pola hubungan antara LPE dengan jumlah industri besar sedang (x_2) terlihat tidak membentuk suatu pola tertentu seperti pada Gambar 2. Oleh sebab itu, metode yang digunakan adalah regresi nonparametrik *spline*.



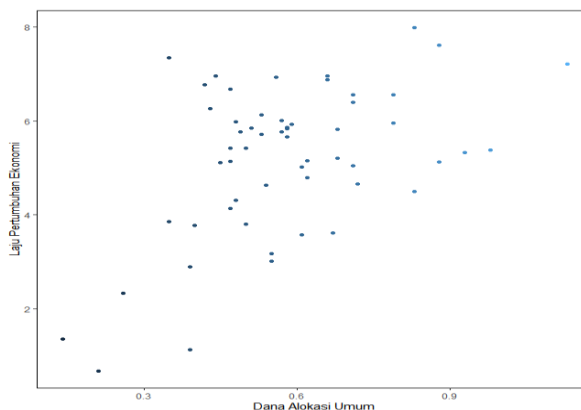
Gambar 3. Scatterplot antara LPE dengan Rata-Rata Lama Sekolah

LPE diduga dipengaruhi oleh rata-rata lama sekolah, dimana rata-rata lama sekolah di Kabupaten/Kota menunjukkan rata-rata tingkat pendidikan yang ditempuh penduduk suatu Kabupaten/Kota. Semakin tinggi tingkat pendidikan suatu penduduk maka peluang untuk memperoleh pekerjaan dan mengurangi tingkat pengangguran juga semakin tinggi. Pola hubungan antara LPE dengan rata-rata lama sekolah (x_3) terlihat tidak membentuk suatu pola tertentu seperti pada Gambar 3. Oleh sebab itu, metode yang digunakan adalah regresi nonparametrik *spline*.



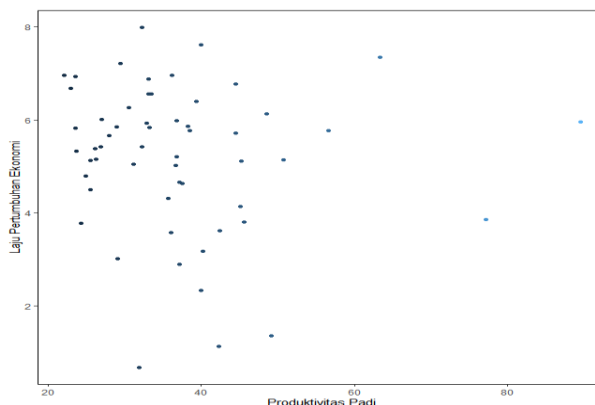
Gambar 4. Scatterplot antara LPE dengan Anggaran Pendapatan dan Belanja Daerah

LPE diduga dipengaruhi oleh APBD, dimana kita ketahui bahwa APBD yang ditujukan untuk pembiayaan proses pembangunan sebagai kegiatan pemerintah dalam meningkatkan kesejahteraan masyarakat. Dengan pemahaman bahwa apabila APBD meningkat maka kemungkinan besar LPE juga akan meningkat. Pola hubungan antara LPE dengan APBD (x_4) terlihat tidak membentuk suatu pola tertentu seperti pada Gambar 4. Oleh sebab itu, metode yang digunakan adalah regresi nonparametrik *spline*.



Gambar 5. Scatterplot antara LPE dengan Dana Alokasi Umum

LPE diduga dipengaruhi oleh dana alokasi umum, dimana kita ketahui bahwa dana alokasi umum sangat penting dalam membantu pemerintah daerah dalam menjaga dan menjamin tercapainya standar pelayanan publik minimum diseluruh negeri yg diharapkan meningkatkan produktivitas kerja dan berpengaruh signifikan terhadap LPE. Pola hubungan antara LPE dengan dana alokasi umum (x_5) terlihat tidak membentuk suatu pola tertentu seperti pada Gambar 5. Oleh sebab itu, metode yang digunakan adalah regresi nonparametrik *spline*.



Gambar 6. Scatterplot antara LPE dengan Produktivitas Padi

LPE diduga dipengaruhi oleh produktivitas padi, ketahanan pangan bagi suatu negara yg memiliki penduduk besar merupakan hal yang sangat penting terutamanya beras yang merupakan makanan pokok orang Indonesia, meningkat atau menurunnya produktivitas padi diharapkan memberikan pengaruh signifikan terhadap LPE. Pola hubungan antara LPE dengan produktivias padi (x_6) terlihat tidak membentuk suatu pola tertentu seperti pada Gambar 6. Oleh sebab itu, metode yang digunakan adalah regresi nonparametrik *spline*.

3.2 Titik Knot Optimal

Tabel 3. Titik Knot Optimal

Variabel Prediktor	Satu Knot	Dua Knot			Tiga Knot	
x_1	73.4	60.6	66.1	60.6	61.1	76.6
x_2	71.6	4,8	33.4	4,8	7.2	88.3
x_3	9	6,1	7.3	6.1	6.2	9.8
x_4	2.3	0,7	1.4	0.7	0.8	2.7
x_5	0.7	0,2	0.4	0.2	0.2	0.9
x_6	63.4	24.8	41.3	24.8	26.2	73
GCV	1.366	1.268			1.208	

Berdasarkan Tabel di atas, perbandingan GCV dari ketiga model pada Tabel 3 maka diperoleh model terbaik yaitu dengan menggunakan *spline* dengan tiga titik knot yaitu dengan nilai GCV sebesar 1.208 Nilai GCV ini adalah yang paling minimum dibandingkan dengan satu knot dan dua knot yang masing-masing nilai GCV-nya adalah 1,366 dan 1,268. Dengan demikian, model *spline* yang digunakan untuk memodelkan data LPE di Pulau Kalimantan adalah model *spline* dengan tiga titik knot. Langkah selanjutnya adalah membuat model regresi nonparametrik dengan tiga titik knot.

$$\hat{y} = 2,567 - 2,935x_1 - 8,871(x_1 - 60,602)_+ + 11,837(x_1 - 61,058)_+ - 0,162(x_1 - 76,559)_+ + 0,364x_2 - 1,006(x_2 - 4,776)_+ + 0,69(x_2 - 7,163)_+ - 0,178(x_2 - 88,347)_+ + 31,346x_3 - 21,761(x_3 - 6,061)_+ - 9,408(x_3 - 6,167)_+ - 1,266(x_3 - 9,761)_+ + 30,439x_4 - 73,961(x_4 - 0,743)_+ + 40,038(x_4 - 0,8)_+ + 9,581(x_4 - 2,722)_+ + 2,896x_5 + 3,004(x_5 - 0,18)_+ + 3,237(x_5 - 0,201)_+ - 35,266(x_5 - 0,888)_+ - 1,22x_6 + 1,904(x_6 - 24,786)_+ - 0,656(x_6 - 26,163)_+ - 0,138(x_6 - 73,007)_+$$

Model regresi nonparametrik *spline* dengan menggunakan tiga titik knot di atas memiliki nilai R^2 sebesar 82.15 persen dan *Mean Square Error* (MSE) sebesar 0,805. Hal ini dapat dikatakan bahwa tingkat partisipasi angkatan kerja, jumlah industri besar dan sedang, rata-rata lama sekolah, anggaran pendapatan dan belanja daerah, dana alokasi umum, dan produktivitas padi berpengaruh sebesar 82.15 persen terhadap laju pertumbuhan ekonomi di Kalimantan sedangkan sisanya sebesar 17,85 persen dipengaruhi oleh variabel lain yang tidak diketahui.

3.3 Pengujian Signifikansi Parameter Model Secara Simultan

Pada pengujian signifikansi parameter model secara simultan digunakan hipotesis sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{24} = 0$$

$$H_1 : \text{minimal terdapat satu } \beta_j \neq 0, j = 1, 2, \dots, 24$$

Hasil ANOVA untuk pengujian parameter model regresi nonparametrik *spline* secara simultan ditunjukkan pada Tabel 4 sebagai berikut:

Tabel 4. ANOVA Model Regresi Spline Secara Simultan

Model	Df	Sum of Square	Mean Square	F _{hitung}	P-value
Regresi	24	114.8	4.8	5.9	3.388 x 10 ⁻⁶
Residual	31	24.9	0.8		
Total	55	139.7			

Berdasarkan hasil ANOVA dari Tabel 4 di atas dapat diketahui bahwa *p-value* adalah sebesar 3.388 x 10⁻⁶, nilai tersebut lebih kecil dari nilai α (0,05). Sehingga berdasarkan hasil tersebut dapat disimpulkan bahwa H_0 ditolak, maka dapat disimpulkan bahwa parameter memiliki pengaruh signifikan terhadap LPE.

3.4 Pengujian Signifikansi Parameter Model Secara Individu

Pada pengujian signifikansi parameter model secara individu digunakan hipotesis sebagai berikut.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, 24$$

Hasil pengujian signifikansi parameter secara individu dapat dilihat pada Tabel 5 sebagai berikut:

Tabel 5. Pengujian Parameter Model Regresi Secara Individu

Variabel	Parameter	Koefisien	t-hitung	P-value
x_1	β_0	2.567	2.656	0.017*
	β_1	-2.935	-1.937	0.062
	β_2	-8.871	-2.347	0.025*
	β_3	11.837	2.8	0.009*
	β_4	-0.162	-0.814	0.422
x_2	β_5	0.364	2.761	0.01*
	β_6	-1.006	-3.199	0.003*
	β_7	0.691	3.246	0.003*
	β_8	-0.178	-3.202	0.003*
x_3	β_9	31.346	2.057	0.048*
	β_{10}	-21.761	-0.897	0.377
	β_{11}	-9.408	-0.672	0.507
	β_{12}	-1.266	-1.509	0.141
x_4	β_{13}	30.439	1.644	0.11
	β_{14}	-73.961	-1.548	0.132
	β_{15}	40.038	1.311	0.199
	β_{16}	9.581	4.221	1.97* x 10⁻⁴
x_5	β_{17}	2.896	7.191	4.38* x 10⁻⁸
	β_{18}	3.003	6.797	1.3* x 10⁻⁷
	β_{19}	3.237	6.574	2.4* x 10⁻⁷
	β_{20}	-35.266	-3.032	4.9* x 10⁻³
	β_{21}	-1.221	-2.425	0.021*
x_6	β_{22}	1.904	1.929	0.063
	β_{23}	-0.656	-1.195	0.24
	β_{24}	-0.138	-1.49	0.146

Berdasarkan Tabel 5 di atas dapat dilihat bahwa terdapat beberapa parameter yang tidak signifikan. Berdasarkan 25 parameter pada model regresi nonparametrik *Spline*, terdapat 11 parameter yang tidak signifikan. Parameter tersebut tidak signifikan pada taraf signifikansi $\alpha(0,05)$. Pengujian parameter dengan taraf signifikansi t hitung > t tabel (2,04) memberikan hasil serupa yaitu terdapat 11 parameter tidak signifikan. Meskipun terdapat 11 parameter yang tidak signifikan, namun secara keseluruhan dapat disimpulkan bahwa keenam variabel berpengaruh terhadap LPE.

3.5 Interpretasi Model

Dari hasil analisis sebelumnya, diperoleh model regresi nonparametrik *Spline* terbaik adalah dengan tiga titik knot. Model terbaik yang diperoleh adalah sebagai berikut:

Apabila variabel x_2, x_3, x_4, x_5 , dan x_6 dianggap konstan, maka besar pengaruh TPAK (x_1) terhadap LPE adalah

$$\hat{y} = 2,567 - 2,935x_1 - 8,871(x_1 - 60,602)_+ + 11,837(x_1 - 61,058)_+ - 0,162(x_1 - 76,559)_+ \quad (15)$$

$$\left\{ \begin{array}{l} 2,567 - 2,935x_1 \\ 540,167 - 11,806x_1 \\ 182,576 - 0,031x_1 \\ 194,979 - 0,193x_1 \end{array} \right\} \begin{array}{l} x_1 < 60,602 \\ 60,602 \leq x_1 < 61,058 \\ 61,058 \leq x_1 < 76,559 \\ x_1 \geq 76,559 \end{array}$$

Pada saat TPAK kurang dari $60,602 \approx 61$ orang dan TPAK bertambah sebanyak 1 orang maka LPE cenderung turun sebesar 2,935 persen. Jika TPAK jumlahnya diantara 60,602 orang dan kurang dari 61,058 orang dan TPAK bertambah sebanyak 1 orang maka LPE cenderung turun sebesar 11,81 persen. Kemudian jika TPAK jumlahnya diantara $61,058 \approx 61$ orang dan $76,559 \approx 77$ orang dan TPAK bertambah sebanyak 1 orang maka LPE cenderung turun sebesar 0,031 persen. Sedangkan jika TPAK jumlahnya lebih dari atau

sama dengan $76,559 \approx 77$ orang dan bertambah sebanyak 1 orang maka LPE cenderung turun sebesar 0,193 persen

4. KESIMPULAN

1. Model regresi nonparametrik *spline* terbaik adalah dengan tiga titik knot. Model ini memiliki nilai GCV paling minimum yaitu sebesar 1,208 dengan R^2 sebesar 82, 15 persen.
2. Hasil pengujian hipotesis model regresi nonparametrik *spline* pada data LPE di Kalimantan ialah pada uji simultan secara keseluruhan semua variabel prediktor berpengaruh terhadap variabel respon dan pada uji individu dari 25 parameter yang terbentuk terdapat 14 parameter yang signifikan dan 11 parameter yang tidak signifikan. Parameter tersebut tidak signifikan karena nilai p -value nya $< \alpha = 0,05$
3. Interpretasi model regresi nonparametrik *spline* pada salah satu variabel yang berpengaruh pada LPE di Kalimantan adalah sebagai berikut:

Pada saat TPAK kurang dari $60,602 \approx 61$ orang dan TPAK bertambah sebanyak 1 orang maka LPE cenderung turun sebesar 2,935 persen. Jika TPAK jumlahnya diantara 60,602 orang dan kurang dari 61,058 orang dan TPAK bertambah sebanyak 1 orang maka LPE cenderung turun sebesar 11,81 persen. Kemudian jika TPAK jumlahnya diantara $61,058 \approx 61$ orang dan $76,559 \approx 77$ orang dan TPAK bertambah sebanyak 1 orang maka LPE cenderung turun sebesar 0,031 persen. Sedangkan jika TPAK jumlahnya lebih dari atau sama dengan $76,559 \approx 77$ orang dan bertambah sebanyak 1 orang maka LPE cenderung turun sebesar 0,193 persen.

DAFTAR PUSTAKA

- [1] D. Arjun, Sifriyani and Syarifuddin, "Analisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka di Kalimantan Menggunakan Regresi Nonparametrik Spline Truncated," in *Prosiding Seminar Nasional Matematika dan Statistika*, Samarinda, 2019.
- [2] Bintariningrum, Merly F and I. Budiantara, "Pemodelan Regresi Nonparametrik Spline Truncated Dan Aplikasinya Pada Angka Kelahiran Kasar Di Surabaya," *Jurnal Sains dan Seni POMITS*, vol. 3, no. 2337-3520, p. 1, 2014.
- [3] E. Montoya, N. Ulloa and V. Miller, "A Simulation Study Comparing Knot Selection Methods With Equally Spaced Knot in a Penalized Regression Spline," *International Journal of Statistics*, vol. 3, no. 3, 2014.
- [4] Badan Pusat Statistik, "Berita Resmi Statistik," *Pertumbuhan Ekonomi Indonesia Triwulan IV-2017*, p. 1, 5 2 2018.
- [5] BPS Provinsi Kalimantan Barat, "Berita Resmi Statistik," *Pertumbuhan Ekonomi Kalimantan Barat Triwulan IV-2017*, p. 2, 5 2 2018.
- [6] BPS Provinsi Kalimantan Tengah, "Berita Resmi Statistik," *Pertumbuhan Ekonomi Kalimantan Tengah Tahun 2017*, p. 2, 5 2 2018.
- [7] BPS Provinsi Kalimantan Utara, "Berita Resmi Statistik," *Pertumbuhan Ekonomi Kalimantan Utara Tahun 2017*, p. 1, 5 2 2018.
- [8] BPS Provinsi Kalimantan Timur, "Berita Resmi Statistik," *Pertumbuhan Ekonomi Kalimantan Timur Tahun 2017*, pp. 1-2, 5 2 2018.
- [9] N. Fitri, Sifriyani and Desi Yunarti, "Nonparametric Geographically Weighted Regression dengan Pendekatan Spline Truncated pada Data Produktivitas Tanaman Padi," in *Prosiding Seminar Nasional Matematika dan Statistika*, 98-105, Samarinda, 2019.
- [10] J. Millah and Sifriyani, "Evaluasi Faktor yang Mempengaruhi Laju Pertumbuhan Ekonomi di Kalimantan Menggunakan Regresi Nonparametrik Spline dengan Pembobot Geografis," in *Prosiding Seminar Nasional Matematika dan Statistika*, Samarinda, 2019.
- [11] I. Polanagau, Sifriyani and Wasono, "Pemodelan Regresi Nonparametrik Spline Truncated dan Aplikasinya Pada Indeks Pembangunan Manusia di Pulau Kalimantan," in *Prosiding Seminar Nasional Matematika dan Statistika*, Samarinda, 2019.
- [12] R. L. Eubank, *Nonparametric Regression and Spline Smoothing*, New York: Marcel Dekker, 1998.
- [13] N. Agustina, Mukid and Moch. Abdul, "Pemodelan Data Indeks Harga Saham Gabungan Menggunakan Regresi Penalized Spline," *Jurnal Gaussian*, vol. 04, no. 03, pp. 603-612, 2015.
- [14] G. Wahba, *Spline Model for Observational Data*, Philadelphia: Siam XII, 1990.
- [15] W. Hardle, *Applied Nonparametric Regression*, Cambridge: Cambridge University Press, 1990.

- [16] I. Budiantara, "Spline dalam Regresi Nonparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang," Institut Teknologi Sepuluh Nopember, Surabaya, 2009.
- [17] Atmanti and Hastriani Dwi, "Investasi Sumber Daya Manusia Melalui Pendidikan," *Jurnal Dinamika Pembangunan*, vol. 2, no. 1, pp. 30-39, 2005.
- [18] Arfan, Nur Aziza and I. Budiantara, "Pendekatan Spline Untuk Estimasi Kurva Regresi Nonparametrik (Studi Kasus Data Angka Kematian Maternal Di Jawa Timur)," *Jurnal Sains dan Seni POMITS*, vol. 3, no. 2337-3520, p. 1, 2014.
- [19] P. Craven and G. Wahba, "Smoothing Noisy Data With Spline Function," *Numerische Mathematik*, vol. 3, no. 1, pp. 2337-3520, 1979.
- [20] D. Wahyudi, Analisa pengaruh penanaman modal asing langsung (FDI) pada kegiatan hulu minyak bumi terhadap output sektor pertambangan dan penggalan regional propinsi (2002-2008), Jakarta: Universitas Indonesia, 2010.
- [21] N. R. Drapper and H. Smith, Analisis Regresi Terapan Edisi Kedua, Jakarta: PT Gramedia Pustaka, 1992.
- [22] D. Gujarati, Dasar-Dasar Ekonometrika, Jakarta: R.C Salemba Empat, 2012.
- [23] D. Gujarati, Econometric Basic, 3rd Edition, Singapura: Mc Grawhill, 1992.
- [24] I. Hasan, Pokok-Pokok Materi Statistika 2, Jakarta: Bumi AKsara, 1999.
- [25] P. Kriesniati, Desi Yuniarti and Darnah A Nohe, "Analisis Korelasi Somers“D Pada Data Tingkat Kenyamanan Siswa-Siswi SMP Plus Melati Samarinda," *Jurnal Barekeng*, vol. 7, no. 2, pp. 31-40, 2013.
- [26] Kusnendi, Ekonomi Sumber Daya Manusia, Jakarta: PPUT, 2003.
- [27] J. E. Lavery, "Shape-preserving, First-derivative-based Parametric and Nonparametric Cubic L1 Spline Curves", *Computer Aided Geometric Design*, vol. 01, no. 23, pp. 31-40, 2006.
- [28] G. Mankin N and dkk, Pengantar Ekonomi Makro, Jakarta: Salemba Empat, 2012.
- [29] Sugiyono, Statistika untuk Penelitian, Bandung: Alfabeta, 2010.
- [30] Suharyadi and Purwanto, Metodologi Penelitian, Jakarta: Gramedia Pustaka Utama, 2004.
- [31] S. Sukirno, Pengantar Teori Makroekonomi, Jakarta: PT Raja Grafindo Persada, 2004.
- [32] M. Sinungan, Produktivitas, Jakarta: Bumi Aksara, 2005.
- [33] Suyanto, "Fraudulent Financial Statement," *Gajah Mada International Journal of Bussiness*, vol. 11, no. 1, pp. 117-144, 2011.
- [34] C. J. Stone, "Nonparametric M-Regression with Free Knot Splines," *Journal of Statistics Planning and Inference*, vol. 130, pp. 183-206, 2004.
- [35] R. E. Walpole, Ilmu Peluang Dan Statistika Untuk Insinyur dan Ilmuwan, Jakarta: Gramedia Pustaka Utama, 1995.
- [36] S. Wignjosoebroto, Pengantar Teknik dan Manajemen Industri, Surabaya: Guna Widya, 2003.

