



Penerapan Cross Validation (CV) dalam Pemilihan Bandwidth Optimal pada Pemodelan Regresi Nonparametrik Kernel (Studi Kasus: Gizi Buruk pada Balita Di Indonesia)

Application of Cross Validation (CV) in Optimal Bandwidth Selection in Kernel Nonparametric Regression Modeling (Case Study: Malnutrition in Toddlers in Indonesia)

Resti Anita Razak, Indah Manfaati Nur, Prizka Rismawati Arum

Universitas Muhammadiyah Semarang, Semarang

Corresponding author: resti.anita18101996@gmail.com, indahmnur@unimus.ac.id, prizka.rismawati@gmail.com

Riwayat Artikel: Dikirim; Diterima; Diterbitkan

Abstrak

Gizi buruk di Indonesia terbilang cukup tinggi, berdasarkan data dari Kementerian Kesehatan (Kemenkes) Indonesia, angka gizi buruk setiap tahun mengalami peningkatan sejak tahun 2016. Maka dari itu perlu suatu analisis untuk melihat faktor-faktor yang mempengaruhi gizi buruk pada balita di Indonesia. Data pada penelitian ini mempunyai pola sebaran data yang tidak membentuk pola tertentu dan melanggar beberapa asumsi sehingga sulit dianalisis dengan analisis parametrik. Maka dari itu metode yang sesuai adalah Regresi Nonparametrik. Regresi Nonparametrik pada penelitian ini menggunakan pendekatan Kernel. Fungsi Kernel yang digunakan dalam penelitian ini yaitu Fungsi Kernel Gaussian dan pemilihan bandwidth optimal dengan menggunakan metode *Cross Validation* (CV). Hasil analisis menunjukkan bandwidth optimal dengan metode CV sebesar 0.25. Adapun nilai R^2 yang dihasilkan metode CV sebesar 84.73%. Nilai MSE yang dihasilkan metode CV sebesar 0.257. karena memiliki nilai R^2 yang cukup besar dan nilai MSE yang kecil, maka dapat disimpulkan metode CV sudah baik dalam pemilihan bandwidth Optimal yang dapat menghasilkan estimasi mendekati data aslinya.

Kata Kunci : Regresi Nonparametrik, Regresi Kernel, Bandwidth, Gizi Buruk

Abstract

Malnutrition in Indonesia was quite high, based on the data from the Indonesian Ministry of Health, the number of malnutrition has increased every year since 2016. Therefore, an analysis was needed to see the factors that influence malnutrition in children under five in Indonesia. The data in this study had a data distribution pattern had no specific pattern and violates several assumptions so that is why this data was difficult to analyze with parametric analysis. Therefore the appropriate method was Nonparametric Regression. Nonparametric regression in this study used the Kernel approach. Kernel function used in this research was Gaussian Kernel Function and optimal bandwidth selection using the Cross Validation (CV) method. The analysis showed that the optimal bandwidth with the CV method is 0.25. The R^2 value generated by the CV method is 84.73%. MSE value generated by the CV method is 0.257. because it has a fairly large R^2 value and a small MSE value, it can be concluded that the CV method is good in choosing Optimal bandwidth that can produce estimates close to the original data.

Keywords: Nonparametric Regression, Kernel Regression, Bandwidth, Malnutrition



PENDAHULUAN

Gizi buruk merupakan salah satu permasalahan yang menjadi perhatian dunia sampai saat ini. Menurut UNICEF (2018), gizi buruk adalah masalah universal yang tidak bisa diabaikan oleh negara di dunia. setiap tahun sekitar 20 juta bayi dilahirkan dengan berat badan kurang.

Indonesia merupakan salah satu yang sedang berjuang mengatasi masalah kesehatan ini. Gizi buruk sangat berbahaya karena mengancam jiwa anak-anak Indonesia yang merupakan generasi penerus bangsa. Menurut data dari Kemenkes Republik Indonesia, menunjukkan sejak tahun 2016 sampai 2018 persentase gizi buruk selalu meningkat. Hal ini di tunjukkan dengan persentase gizi buruk di Indonesia pada tahun 2016 sebesar 3.4%, pada tahun 2017 sebesar 3.80% dan pada tahun 2018 sebesar 3.9%.

Ada beberapa faktor yang mempengaruhi kejadian gizi buruk, seperti pemberian Asi Eksklusif, Persentase Pemberian Tablet Tambah Darah (Fe) dan Persentase Penduduk Miskin. Maka perlu dilakukan penelitian untuk mengetahui faktor-faktor yang mempengaruhi kejadian gizi buruk di Indonesia. Analisis regresi adalah suatu metode untuk melihat bagaimana pengaruh variabel independen terhadap variabel dependen yang membentuk suatu model. Namun data pada penelitian ini belum ditemukan pola sebaran datanya sehingga diterapkan analisis Regresi Nonparametrik. Penelitian ini menggunakan pendekatan kernel. Pendekatan kernel mempunyai bentuk yang fleksibel dan secara matematik mudah dikerjakan serta mempunyai rata-rata kekonvergenan yang relative cepat (Nisa, 2016).

Penelitian ini menggunakan fungsi Kernel Gaussian. Fungsi kernel ini digunakan karena fungsi kernel Gaussian lebih mudah dalam perhitungan dan penggunaannya serta lebih sering digunakan sedangkan fungsi kernel lainnya memerlukan syarat dalam pengerjaannya. Pada regresi kernel dikenal salah satu estimator yang biasa digunakan untuk mengestimasi suatu fungsi regresi yaitu estimator kernel Nadaraya Watson. yaitu Estimasi dengan pendekatan ini bergantung pada dua parameter yaitu bandwidth (pemulus) dan fungsi kernel. Bandwidth (pemulus) merupakan pengontrol kemulusan kurva.

Pemilihan bandwidth yang optimal untuk mendapatkan kurva regresi yang optimal. Salah satu teknik untuk mendapatkan bandwidth optimal yaitu Cross Validation (CV). Metode ini lebih mudah secara matematisnya. Selain itu metode CV sesuai digunakan untuk data yang memiliki bobot tiap unit observasi yang berbeda setiap unit observasi.

Penelitian-penelitian sebelumnya dengan pemodelan spasial telah banyak dilakukan seperti penelitian Nanda (2016) tentang analisis pengaruh jumlah uang beredar dan nilai tukar rupiah terhadap indeks harga saham gabungan menggunakan pemodelan regresi semiparametrik kernel. Dewi (2017) tentang Pemodelan Jumlah Penduduk Miskin dengan Pendekatan Regresi Nonparametrik Estimator Kernel Nadaraya-Watson dan Penelitian mengenai gizi buruk dalam penelitian Abdul Karim dan Rochdi Wasono (2016) tentang *Modelling Malnutrition Toddlers in East Java Province Using Spatial Regression*, serta penelitian Selvi (2017) tentang Identifikasi factor-faktor yang mempengaruhi gizi buruk dan gizi kurang pada balita menggunakan regresi spasial. Melalui penelitian ini peneliti mencoba mengembangkan metode yang telah digunakan dalam penelitian terdahulu. Sehingga peneliti mengangkat judul. Penerapan Cross Validation (CV) dalam Pemilihan Bandwidth Optimal pada Pemodelan Regresi Nonparametrik Kernel Gaussian (Studi Kasus : Gizi Buruk pada Balita di Indonesia).



METODE

1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Profil Kesehatan Indonesia Tahun 2018, berupa 34 provinsi di Indonesia sebagai unit observasi.

2. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari variabel dependen (Y) dan variabel independen (X).

Y : persentase gizi buruk di Indonesia tahun 2018

X1 : persentase pemberian ASI Eksklusif,

X2 : persentase pemberian tablet tambah darah (Fe)

X3 : persentase penduduk miskin.

3. Analisis Data

Tahapan analisis yang digunakan dalam penelitian ini adalah :

- Membuat analisis statistika deskriptif
- Menguji Asumsi Klasik
- Melihat pola sebaran data yang akan digunakan dalam penelitian dengan cara membuat scatterplot pada setiap variabel.
- Menentukan fungsi kernel yang digunakan, yaitu kernel Gaussian

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right),$$

$$-\infty < x < \infty$$

- Melakukan pemilihan bandwidth (h) yang optimal dan Memodelkan nilai bandwidth (h) optimal dengan metode Cross Validation (CV) yang minimum

$$CV = n^{-1} \sum_{j=1}^n \left[\frac{Y_j - \hat{m}_h(X_j)}{1 - H_{jj}(h)} \right]^2$$

- Mengestimasi dengan estimator Nadaraya- Watson

$$\hat{m}(x) = \frac{\sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) Y_i}{\sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)}$$

- Mengestimasi model Nonparametrik

$$y_i = f(x_i) + \varepsilon_i \quad i = 1, 2, \dots, n$$

Keterangan

y_i : variabel prediktor ke - i

$f(x_i)$: fungsi nonparametrik ke- I yang tidak diketahui

ε_i : error ke-i yang diasumsikan menyebar $N \sim (0; \sigma^2)$

- Menghitung nilai Mean Square Error (MSE) dan koefisien determinasi (R^2) untuk melihat seberapa baik metode CV digunakan dalam pemilihan bandwidth optimal

$$MSE(h) = N^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$R^2 = \frac{JKR}{JKT} = \frac{JKR}{JKR + JKG}$$



Dengan :

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

- i. Mengambil kesimpulan yang di dapatkan dari hasil penelitian.

ANALISIS DAN PEMBAHASAN

1. Statistika Deskriptif Data

Tabel 1 :

Hasil Deskriptif Data

	Min	Max	Mean	SD	Varians
Y	2	7.4	4.39	1.54	2.37
X1	20.43	80.28	61.53	13.39	179.40
X2	32.11	99.49	77.37	16.58	274.85
X3	3.55	27.43	10.61	5.70	32.53

Tabel 1 dapat diinterpretasikan hasil statistik deskriptif dari masing-masing variabel.

Berdasarkan hasil deskriptif data, terlihat bahwa persentase gizi buruk di Indonesia terendah sebesar 2% dan tertinggi sebesar 7.4%. Persentase pemberian ASI eksklusif di Indonesia terendah sebesar 20.43% dan tertinggi sebesar 80.28%. Persentase pemberian tablet tambah darah (Fe) di Indonesia terendah sebesar 32.11% dan tertinggi sebesar 99.49%. Persentase penduduk miskin di Indonesia terendah sebesar 3.55% dan tertinggi sebesar 27.43%.

2. Uji Asumsi Klasik

a. Kenormalan

Uji kenormalan menggunakan uji Anderson Darling, dengan hasil sebagai berikut,

Tabel 2 :

Hasil Uji Kenormalan

Asumsi	p-value	Keterangan
Kenormalan	3.534e-05	Terjadi Autokorelasi

Berdasarkan hasil uji autokorelasi dengan uji Durbin Watson yang terlihat pada tabel 4.2, terlihat bahwa dalam data penelitian ini terjadi autokorelasi. Hal ini dibuktikan dengan p-value < (0.05).

Uji asumsi selanjutnya yaitu uji autokorelasi. Pengujian autokorelasi dalam penelitian ini menggunakan uji Durbin Watson, dengan hasil sebagai berikut:

Tabel 3 :

Hasil Uji Autokorelasi

Asumsi	p-value	Keterangan
Autokorelasi	3.534e-05	Terjadi Autokorelasi

Berdasarkan hasil uji autokorelasi dengan uji Durbin Watson yang terlihat pada tabel 3, terlihat bahwa dalam data penelitian ini terjadi autokorelasi. Hal ini dibuktikan dengan p-value < α (0.05).

b. Homogenitas



Pengujian Heteroskedasitas dalam penelitian ini menggunakan uji Breusch-Pagan, dengan hasil sebagai berikut :

Tabel 4 :
Hasil Uji Heteroskedasitas

Asumsi	p-value	Keterangan
Heteroskedasitas	0.07749	Tidak Terjadi Heteroskedasitas

Berdasarkan hasil uji autokorelasi dengan uji Breusch-Pagan yang terlihat pada tabel 4, terlihat bahwa dalam data penelitian ini tidak terjadi Heteroskedasitas. Hal ini dibuktikan dengan p-value > α (0.05).

c. Multikolinearitas

Uji asumsi berikutnya yaitu uji Multikolonearitas. Pengujian Multikolinearitas dalam penelitian ini menggunakan uji *Varince Inflation Factor* (VIF), dengan hasil sebagai berikut :

Tabel 5 :
Hasil Uji VIF

Variable	VIF	Keterangan
X1	1.24	Tidak terjadi multikolinearitas
X2	1.38	Tidak terjadi multikolinearitas
X3	1.22	Tidak terjadi multikolinearitas

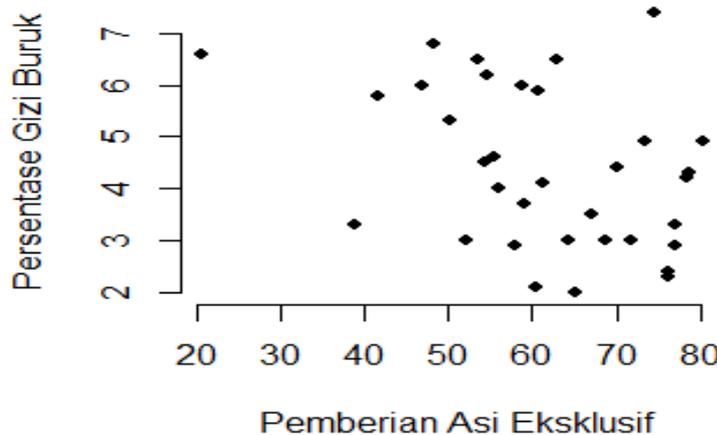
Tabel 5 menunjukkan nilai VIF setiap variabel predictor dalam penelitian ini. Terlihat bahwa setiap variable memiliki nilai vif yang lebih kecil dari 10 maka dapat disimpulkan tidak terjadi multikolinearitas dalam data penelitian ini.

Regresi Non Parametrik Kernel

a. Identifikasi Pola Sebaran Data

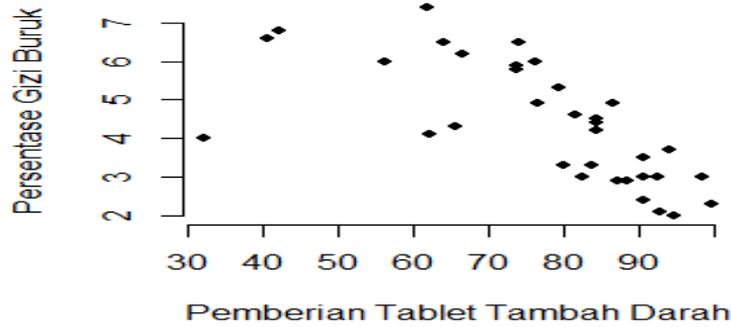
Gambar 1 :

Scatterplot dari X1 terhadap Y

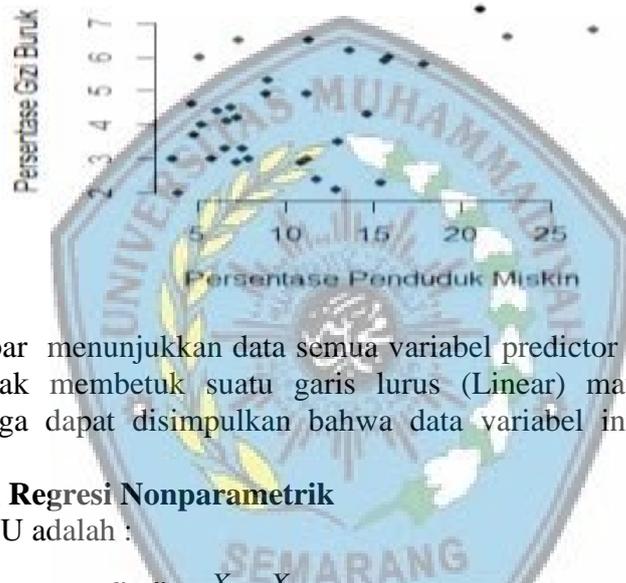




Gambar 2 :
Scatterplot dari X2 terhadap Y



Gambar 3 :
Scatterplot dari X3 terhadap Y



Ketiga gambar menunjukkan data semua variabel predictor terhadap variabel respon menyebar acak tidak membentuk suatu garis lurus (Linear) maupun suatu lengkungan (Nonlinear), sehingga dapat disimpulkan bahwa data variabel ini merupakan komponen nonparametrik.

b. Estimasi Model Regresi Nonparametrik

Dengan nilai U adalah :

$$u = \frac{x - x_i}{h} = \frac{X_i - X_j}{h} \quad \text{dengan } i=1,2,3,\dots,34$$

$$j = 1,2,3,\dots,34$$

Diperoleh fungsi kernel yang sesuai untuk penelitian ini adalah Fungsi Kernel Gaussian sebagai berikut :

$$K_h(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{X_j - X_i}{h}\right)^2\right)$$

Dengan menggunakan variabel nonparametrik yaitu pemberian asi eksklusif, persentase pemberian tablet tambah darah dan persentase penduduk miskin, estimasi diperoleh dengan mensubstitusikan fungsi kernel ke estimator Nadaraya Watson, sebagai berikut :



$$\hat{m}(x) = \frac{\sum_{i=1}^{34} \left(\frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{X_j - X_i}{h} \right)^2 \right) Y_i \right)}{\sum_{i=1}^{34} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{X_j - X_i}{h} \right)^2 \right)} + \varepsilon_i$$

sehingga didapatkan model berikut :

$$y = m(x) + \varepsilon_i$$

$$y = \frac{\sum_{i=1}^{34} \left(\frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{X_j - X_i}{h} \right)^2 \right) Y_i \right)}{\sum_{i=1}^{34} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{X_j - X_i}{h} \right)^2 \right)} + \varepsilon_i$$

3. Pemilihan Bandwidth Optimal

Kemudian langkah selanjutnya mencari bandwidth optimum, CV merupakan metode yang digunakan dalam menentukan pemilihan bandwidth optimal, dilakukan dengan memilih nilai CV yang paling minim. Diperoleh lah nilai bandwidth dan CV minumumnya :

Tabel 6 :
Nilai Bandwidth beserta nilai CV

No	Bandwidth	CV
1	0.25	0.397
2	0.26	0.412
3	0.27	0.426
4	0.28	0.440
5	0.29	0.454

Berdasarkan hasil diatas terlihat bahwa nilai bandwidth optimal menurut CV sebesar 0.25.

Berikut persamaan model persentase gizi buruk dengan pendekatan nonparametrik kernel *Nadaraya-Watson* dengan metode pemilihan bandwidth optimal CV adalah sebagai berikut :

$$\hat{y} = \frac{\sum_{p=1}^{34} \left(\frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_1 - X_{1p}}{0.25} \right)^2 \right) Y_i \right)}{\sum_{p=1}^{34} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_1 - X_{1p}}{0.25} \right)^2 \right)} + \frac{\sum_{p=1}^{34} \left(\frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_2 - X_{2p}}{0.25} \right)^2 \right) Y_i \right)}{\sum_{p=1}^{34} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_2 - X_{2p}}{0.25} \right)^2 \right)} + \frac{\sum_{p=1}^{34} \left(\frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_3 - X_{3p}}{0.25} \right)^2 \right) Y_i \right)}{\sum_{p=1}^{34} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_3 - X_{3p}}{0.25} \right)^2 \right)}$$



4. Pemilihan Metode yang Baik

Setelah itu diperoleh lah nilai R^2 dan MSE dengan hasil sebagai berikut :

Tabel 4.8

	R^2	MSE
CV	84.73%	0.257

Tabel 4.8 menunjukkan nilai koefisien Determinasi dan MSE dari metode CV. Berdasarkan tabel 4.8, nilai R-sq dan MSE terlihat bahwa CV merupakan metode yang baik untuk menentukan bandwidth optimal dalam pemodelan Regresi Nonparametrik Kernel Gaussian pada penelitian ini, karena metode CV menghasilkan nilai R^2 yang besar dan MSE yang kecil.

KESIMPULAN

Berdasarkan hasil penelitian metode CV merupakan metode yang baik dalam pemilihan bandwidth optimal dalam kasus kejadian gizi buruk pada balita di Indonesia karena memiliki nilai R^2 yang besar dan nilai MSE yang kecil. Dengan nilai bandwidth sebesar 0.25 dan nilai CV sebesar 0.397, maka diperoleh persamaan sebagai berikut :

$$\hat{y} = \frac{\sum_{p=1}^{34} \left(\frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_1 - X_{1p}}{0.25} \right)^2 \right) \right) Y_i}{\sum_{p=1}^{34} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_1 - X_{1p}}{0.25} \right)^2 \right)} + \frac{\sum_{p=1}^{34} \left(\frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_2 - X_{2p}}{0.25} \right)^2 \right) \right) Y_i}{\sum_{p=1}^{34} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_2 - X_{2p}}{0.25} \right)^2 \right)}$$

$$+ \frac{\sum_{p=1}^{34} \left(\frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_3 - X_{3p}}{0.25} \right)^2 \right) \right) Y_i}{\sum_{p=1}^{34} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_3 - X_{3p}}{0.25} \right)^2 \right)}$$

DAFTAR PUSTAKA

- Adzim, Muhammad Abdul. 2015. *Fungsi Kernel Gaussian untuk memodelkan data UAN SMA Al'Mahadul Islami Beji Bangil Pasuruan*. Skripsi. UIN Maulana Malik Ibrahim Malang
- Ayuningtyas, Tri. 2018. *Regresi Nonparametrik Kernel Nadaraya Watson dalam Data Time Series*. Skripsi. UII Yogyakarta
- Badan Pusat Statistik (BPS) Indonesia . 2018. *Konsumsi Kalori dan Protein Penduduk Indonesia dan Provinsi, September 2018*.
- Devi, Alvita Rachma. 2018. *Metode Unbiased Risk (UBR) dan Cross Validation (CV) Untuk Pemilihan Titik Knot Optimal Dalam Regresi Nonparametrik Multivariabel Spline Truncated (Studi Kasus : Data Tingkat Pengangguran Terbuka di Provinsi Jawa Tengah tahun 2015)*. Skripsi. ITS



- Dewi, Nadya Permata Tungga. 2018. *Pemodelan Jumlah Penduduk Miskin dengan Pendekatan Regresi Nonparametrik Estimator Kernel Nadaraya-Watson*. Skripsi. UNIMUS
- Esty. 2014. *Regresi Kernel dengan Nadaraya Watson*. Skripsi. UNY
- Fitriani, Anna, et al. 2015. *Estimasi Model Regresi Semiparametrik Menggunakan Estimator Kernel Uniform*. Jurnal Matematika. Vol. 4 No 4. Universitas Udayana
- Hardle, W. 1990. *Smoothing Techniques With Implementation in S*. Belgium: Universite Catholique de Louvain
- Indrayanti, Anisa Ika. 2016. *Estimator Kernel Cosinus dan Kernel Gaussian dalam Model Regresi Nonparametrik pada Data Butterfly Diagram Siklus Aktivitas Matahari ke-23 (Studi Kasus di BPD LAPAN Watukosek)*. Skripsi. UIN Maulana Malik Ibrahim Malang
- Irharni. 2017. *Pengaruh Jumlah Penduduk, Penangguran, dan Pengeluaran Pemerintah Terhadap Kemiskinan di Indonesia tahun 1986-2015*. Skripsi. UNY
- Junaidi. 2015. *Statistika Non Parametrik*. Research Gate. Universitas Jambi
- Karim, A. and Wasono, R., 2016. *Modelling Malnutrition Toddlers in East Java Province Using Spatial Regression*. Artikel Ilmiah
- Kementerian Kesehatan Indonesia. 2013. *Infodatin pusat data dan informasi kementerian RI - Situasi dan Analisis Gizi*
- Kementerian Kesehatan Indonesia. 2017. *WartaKesmas- Gizi, Investasi Masa Depan Bangsa (Edisi 2)*. Wartakesmas
- Kementerian Kesehatan Indonesia. 2016-2018. *Profil Kesehatan Indonesia*
- Kurniasih, Dedeh, et al. 2013. *Efisiensi Relatif Estimator Fungsi Kernel Gaussian. Terhadap Estimator Polinomial Dalam Peramalan USD Terhadap JPY*. UNNES Journal of Mathematics. Vol. 2 No. 2. UNNES
- Marisa, Rifada, et al. 2013. *Permodelan kejadian gizi buruk pada balita di Surabaya berdasarkan pendekatan regresi spasial*. Prosiding. UNY
- Malin, Adella Fitria. 2016. *Penduga Regresi Nonparametrik dan Nonlinear dengan Metode Priestley-Chao, Nadaraya Watson dan Metode Fourier*. Skripsi. Unlam
- Nanda, Deden Aditya, 2016. *Analisis Pengaruh Jumlah Uang Beredar dan Nilai Tukar Rupiah Terhadap Indeks Harga Saham Gabungan Menggunakan Pemodelan Regresi Semiparametrik Kernel*. Jurnal Gaussian. Vol. 5 no. 3. Undip
- Nisa, Sofiyatin. 2016. *Estimator Kernel Apechnikov dan Kernel Triangel pada data rata-rata bulanan Bilangan Sunspot, NOAA*. Skripsi. UIN Maulana Malik Ibrahim Malang
- Suparti, et al. 2017. *Regresi Nonparametrik*. Ponorogo: WADEgroup
- United Nations Emergency Children's Fund (UNICEF). *Nutrisi*. <https://www.unicef.org>. (diakses tanggal 11 Mei 2019)
- Utami, Endah Suryaningsih. 2018. *Pemodelan Produksi Padi di Jawa Timur Menggunakan Regresi Semiparametrik Kernel*. Skripsi. Unimus
- Yuniartika, Yuyun, et al. 2013. *Penentuan Generalized Cross Validation (GCV) Sebagai Kriteria Dalam Pemilihan Model Regresi B-Spline Terbaik*. Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster). Vol.2 No. 2, hal 121-126. Untan