



**PEMODELAN REGRESI NONPARAMETRIK
ESTIMATOR SPLINE DAN KERNEL
STUDI KASUS: PENGARUH RATA-RATA LAMA SEKOLAH TERHADAP
PERSENTASE PENDUDUK MISKIN DI KALIMANTAN TAHUN 2023**

SKRIPSI

**Untuk memenuhi persyaratan
Dalam menyelesaikan program sarjana Strata-1 Statistika**

**Oleh
MUHAMMAD FIRDAUS
NIM. 2011017210013**

**PROGRAM STUDI S-1 STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
MEI 2024**

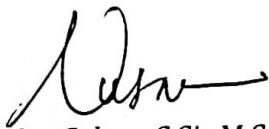
SKRIPSI

Pemodelan Regresi Nonparametrik Estimator Spline Dan Kernel Studi Kasus: Pengaruh Rata-rata Lama Sekolah Terhadap Persentase Penduduk Miskin di Kalimantan Tahun 2023

Oleh:
Muhammad Firdaus
2011017210013

Telah dipertahankan pada hari Rabu, tanggal 22-05-2024 dan telah disetujui oleh dosen pembimbing dan dosen penguji sebagai berikut:

Pembimbing I



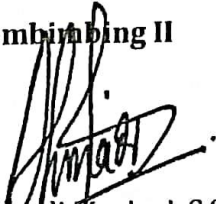
Nur Salam, S.Si., M.Sc.
NIP. 197708132005011003

Penguji I



Prof. Dewi Anggraini, S.Si., M.App.Sci., Ph.D.
NIP. 198303282005012001

Pembimbing II



Ahmad Murjani, S.ST., M.PP., Ph.D.
NIP. 198207242004121003

Penguji II



Selvi Annisa, S.Si., M.Si.
NIP. 199212262022032016

Banjarbaru, 13 Juni 2024

Mengetahui,

Koordinator Program Studi Statistika
MIPA ULM




Prof. Dewi Anggraini, S.Si., M.App.Sci., Ph.D.
NIP. 198303282005012001

PERNYATAAN

Dengan ini Saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan Saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 22 Mei 2024



18B0DAKX807186103 mad Firdaus

NIM. 2011017210013

ABSTRAK

**PEMODELAN REGRESI NONPARAMETRIK ESTIMATOR SPLINE DAN KERNEL
STUDI KASUS: PENGARUH RATA-RATA LAMA SEKOLAH TERHADAP
PERSENTASE PENDUDUK MISKIN DI KALIMANTAN TAHUN 2023** (Oleh:
Muhammad Firdaus; Pembimbing: Nur Salam, S.Si., M.Sc.; Ahmadi Murjani, S.ST., M.PP.,
Ph.D. ; 84 halaman)

Kemiskinan merupakan masalah yang dapat disebabkan karena rendahnya kualitas maupun kuantitas pendidikan. Namun, pola hubungan antara data persentase penduduk miskin dengan rata-rata lama sekolah belum dapat ditentukan secara jelas karena tidak membentuk pola tertentu. Oleh karena itu, pemodelan regresi nonparametrik akan lebih relevan pada kasus data tersebut. Penelitian ini bertujuan untuk menggambarkan pengaruh rata-rata lama sekolah terhadap persentase penduduk miskin dengan model terbaik antara regresi spline dan kernel. Regresi yang digunakan pada penelitian ini yaitu regresi spline *truncated* dan regresi kernel *Nadaraya-Watson Estimator* dengan fungsi *Gaussian*, *Epanechnikov*, dan *Triangular*. Beberapa metode tersebut menghasilkan performa yang berbeda sehingga perlunya penentuan model yang optimal. Hasil pemilihan model menunjukkan bahwa regresi spline *truncated* 3 knot merupakan model yang dipilih pada data tersebut berdasarkan R^2 tertinggi sebesar 0,298; MSE sebesar 3,159; dan MAPE sebesar 26,615. Hasil pengujian model terpilih, diperoleh bahwa variabel rata-rata lama sekolah berpengaruh terhadap tingkat persentase penduduk miskin. Semakin tinggi nilai rata-rata lama sekolah, semakin rendah persentase penduduk miskin.

Kata Kunci : Pendidikan, Kemiskinan, Regresi, Spline, Kernel

ABSTRACT

NONPARAMETRIC REGRESSION MODELING SPLINE AND KERNEL ESTIMATORS CASE STUDY: THE INFLUENCE OF MEAN YEAR OF SCHOOLING ON THE PERCENTAGE OF THE POOR POPULATION IN KALIMANTAN IN 2023 (By: Muhammad Firdaus; Supervisors: Nur Salam, S.Si., M.Si.; Ahmadi Murjani, S. ST., M.PP., Ph.D.; 84 pages)

Poverty is a problem that can be caused by low quality and quantity of education. However, the pattern of relationship between data on the percentage of poor people and mean year of schooling cannot be clearly determined because it does not form a particular pattern. Therefore, nonparametric regression modeling will be more relevant in the case of this data. This research aims to describe the effect of average years of schooling on the percentage of poor people using the best model between spline and kernel regression. The regression used in this research is truncated spline regression and Nadaraya-Watson Estimator kernel regression with Gaussian, Epanechnikov and Triangular functions. Some of these methods produce different performance so it is necessary to determine the optimal model. The model selection results show that 3 knot truncated spline regression is the model chosen for the data based on the highest Rsquare of 0.298; MSE of 3.159; and MAPE of 26.615. The results of testing the selected model showed that the variable mean year of schooling had an effect on the percentage level of the poor population. The higher the mean year of schooling, the lower the percentage of poor people.

Keywords: Education, Poverty, Regression, Spline, Kernel

PRAKATA

Alhamdulillahirobbilalamin, puji syukur saya panjatkan ke hadirat Allah swt. atas segala rahmat dan karunia yang telah dilimpahkan, serta izin-Nya kepada saya sehingga dapat menyelesaikan Tugas Akhir atau Skripsi saya yang merupakan syarat untuk mendapatkan gelar sarjana. Proyek tugas akhir ini saya laksanakan mulai tanggal 27 Agustus 2023 dengan memperhatikan kasus, berita, maupun data yang membuat saya tertarik untuk menjadikannya sebagai topik tugas akhir.

Dalam penyusunan laporan ini, saya mengucapkan terimakasih kepada pihak-pihak yang telah berkontribusi baik secara langsung maupun tidak langsung, dari awal hingga selesainya skripsi ini. Diantaranya:

1. Allah SWT, yang selalu memberikan saya kekuatan dan kelancaran dalam penyelesaian tugas.
2. Bapak Nur Salam, selaku pembimbing I yang telah mendampingi saya dalam penyusunan dan memberikan kritik, saran, dan referensi untuk metode dalam proposal saya.
3. Bapak Ahmadi Murjani, selaku pembimbing II yang telah membimbing saya dan memberikan pemahaman terhadap analisis maupun data yang saya gunakan dalam proposal ini.
4. Ibu Dewi Anggraini, selaku penguji I saya dalam seminar proposal tugas akhir yang telah memberikan penilaian dan masukan kepada saya.
5. Ibu Selvi Annisa, selaku penguji II saya dalam seminar proposal tugas akhir yang telah memberikan penilaian dan masukan kepada saya.
6. Ibu Maisarah, selaku dosen pembimbing akademik yang telah memberikan arahan dan bimbingan akademik selama masa perkuliahan.
7. Orang Tua, yang selalu mendukung dan mendoakan yang terbaik untuk saya dalam memperoleh gelar S1.
8. Kawan-kawan PURWAKARTA, yang selalu memberi motivasi dan saling mendukung untuk lulus sama-sama.
9. BEM FMIPA & HIMASTA "INFINITE" yang mewadahi saya untuk berproses selama menjadi mahasiswa.
10. Teman-teman yang bersangkutan, yang telah memberikan masukan, pendapat, serta dukungan untuk proposal ini.

Dalam penyusunan skripsi ini, saya menyadari bahwa masih banyak terdapat kekurangan baik pada teknis penulisan, materi, maupun dalam kelengkapan informasi. Untuk itu kritik dan saran membangun dari semua pihak sangat saya harapkan demi penyempurnaan penyusunan skripsi ini. Akhir kata penulis berharap semoga skripsi ini bermanfaat bagi yang lainnya.

DAFTAR ISI

PERNYATAAN	iii
PRAKATA.....	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
DAFTAR ISTILAH.....	xi
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	3
1.4. Manfaat Penelitian	3
1.4.1. Manfaat Teoritis	3
1.4.2. Manfaat Praktis	3
1.5. Batasan Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Kajian Penelitian Terdahulu	5
2.2. Kajian Teori	6
2.2.1. Kemiskinan.....	6
2.2.2. Rata-rata Lama Sekolah	8
2.2.3. Statistika Deskriptif	10
2.2.4. Analisis Regresi.....	10
2.2.5. Regresi Nonparametrik.....	11
2.2.6. Regresi Nonparametrik Spline	12
2.2.7. Estimator Parameter Spline <i>Truncated</i>	12
2.2.8. Titik <i>Knot</i>	13
2.2.9. Regresi Nonparametrik Kernel	14
2.2.10. Estimator Parameter Kernel <i>Nadaraya-Watson</i>	14
2.2.11. Fungsi Kernel.....	15
2.2.12. <i>Bandwidth</i>	16
2.2.13. <i>General Cross Validation</i>	17
2.2.14. Uji Parameter Model.....	18
2.2.15. Pengukuran Keباikan dan Akurasi Model	19
BAB III METODE PENELITIAN.....	20

3.1.	Sumber Data	20
3.2.	Variabel Penelitian	20
3.3.	Prosedur Penelitian.....	21
BAB IV HASIL DAN PEMBAHASAN		24
4.1.	Identifikasi Data Penelitian.....	24
4.2.	Statistika Deskriptif.....	28
4.3.	<i>Scatter plot</i>	30
4.4.	Pemodelan Spline <i>Truncated</i>	30
4.3.1.	Pemilihan Titik <i>Knot</i> Optimal Dengan Satu Titik <i>Knot</i>	30
4.3.2.	Pemilihan Titik <i>Knot</i> Optimal Dengan Dua Titik <i>Knot</i>	31
4.3.3.	Pemilihan Titik <i>Knot</i> Optimal Dengan Tiga Titik <i>Knot</i>	32
4.3.4.	Pemilihan Titik <i>Knot</i> Terbaik.....	34
4.3.5.	Estimasi Parameter Model Spline Terpilih.....	34
4.5.	Pemodelan Kernel	34
4.3.1.	Pemilihan <i>Bandwidth</i> Optimal Kernel <i>Gaussian</i>	34
4.3.2.	Pemilihan <i>Bandwidth</i> Optimal Kernel <i>Epanechnikov</i>	35
4.3.3.	Pemilihan <i>Bandwidth</i> Optimal Kernel <i>Triangular</i>	36
4.3.4.	Pemilihan Fungsi Kernel Optimal	36
4.3.5.	Estimasi Parameter Model Kernel Terpilih.....	37
4.6.	Pemilihan Model Terbaik	37
4.7.	Pengujian Signifikansi Parameter Model Terpilih	38
4.8.	Interpretasi Model Terpilih Berdasarkan Studi Kasus	39
BAB V PENUTUP.....		42
5.1.	Kesimpulan.....	42
5.2.	Saran.....	42
DAFTAR PUSTAKA.....		43
LAMPIRAN.....		46
RIWAYAT HIDUP.....		72

DAFTAR GAMBAR

Gambar 1 Persentase Penduduk Miskin di Indonesia Tahun 2023	1
Gambar 2 Rata-rata Lama Sekolah di Indonesia Tahun 2023	2
Gambar 3 Diagram Alir Alur Penelitian.....	23
Gambar 4 Grafik RLS dan PPM Prov. Kalimantan Barat.....	24
Gambar 5 Grafik RLS dan PPM Prov. Kalimantan Selatan.....	25
Gambar 6 Grafik RLS dan PPM Prov. Kalimantan Tengah	26
Gambar 7 Grafik RLS dan PPM Prov. Kalimantan Utara.....	27
Gambar 8 Grafik RLS dan PPM Prov. Kalimantan Timur.....	28
Gambar 9 Scatter plot RLS dengan PPM.....	30
Gambar 10 Visualisasi Spline Satu Titik Knot.....	31
Gambar 11 Visualisasi Spline Dua Titik Knot.....	32
Gambar 12 Visualisasi Spline Tiga Titik Knot.....	33
Gambar 13 Visualisasi Kernel Gaussian	35
Gambar 14 Visualisasi Kernel Epanechnikov.....	35
Gambar 15 Visualisasi Kernel Triangular	36
Gambar 16 Visualisasi Plot Spline Truncated 3 Titik Knot	39

DAFTAR TABEL

Tabel 1 Jenjang Pendidikan dan Nilai Konversi.....	9
Tabel 2 Tipe Pola Data	11
Tabel 3 Daftar Fungsi Kernel	15
Tabel 4 ANOVA Uji Parameter.....	18
Tabel 5 Sumber Data Penelitian	20
Tabel 6 Variabel Penelitian.....	21
Tabel 7 Analisis Deskriptif Variabel	29
Tabel 8 Pemilihan Titik Knot Optimal dengan 1 Knot.....	30
Tabel 9 Pemilihan Titik Knot Optimal dengan 2 Knot.....	31
Tabel 10 Pemilihan Titik Knot Optimal dengan 3 Knot	33
Tabel 11 Perbandingan Nilai GCV.....	34
Tabel 12 Pemilihan Bandwidth Optimal Kernel Gaussian	34
Tabel 13 Pemilihan Bandwidth Optimal Kernel Epanechnikov	35
Tabel 14 Pemilihan Bandwidth Optimal Kernel Triangular	36
Tabel 15 Pemilihan Fungsi Kernel Optimal.....	37
Tabel 16 Pemilihan Model Terbaik.....	37
Tabel 17 Tabel ANOVA	38
Tabel 18 Tabel Uji Parsial	39

DAFTAR ISTILAH

<i>ANOVA</i>	: <i>Analysis of Variance</i>
<i>BPS</i>	: Badan Pusat Statistik
<i>MSE</i>	: <i>Mean Square Error</i>
<i>MAPE</i>	: <i>Mean Absolute Percentage Error</i>
<i>GCV</i>	: <i>General Cross Validation</i>
<i>OLS</i>	: <i>Ordinary Least Square</i>
<i>LPE</i>	: <i>Local Polynomial Estimator</i>
<i>NWE</i>	: <i>Nadaraya-Watson Estimator</i>
<i>WLS</i>	: <i>Weighted Least Square</i>
R^2	: <i>R Square</i> (Koefisien determinasi)
α	: Intersep
β	: Kemiringan/ <i>Gradien/Slope</i>
Y_i	: Variabel dependen
x_i	: Variabel independen
$f(x_i)$: Kurva regresi yang tidak diketahui bentuknya
ε_i	: Galat/ <i>Error</i>
β_j	: Parameter model polinomial, $j = 1, 2, \dots, m$.
$\beta_{(k+m)}$: Parameter komponen <i>truncated</i> , $k = 1, 2, \dots, r$.
λ_k	: Titik-titik <i>knot</i>
r	: Banyak <i>knot</i>
tr	: <i>Trace</i> matriks
\mathbf{X}'	: Matriks <i>transpose</i>
\hat{y}_i	: Taksiran variabel dependen pada observasi ke- i
y_i	: Variabel dependen pada observasi ke- i
n	: Banyak data
h	: <i>Bandwidth</i>
\mathbf{I}	: Matriks identitas
$\%PM_P$: Persentase penduduk miskin di provinsi P
$\%PM_K$: Persentase penduduk miskin di kota/kabupaten K
PM_P	: Jumlah penduduk miskin di provinsi P
PM_K	: Jumlah penduduk miskin di kota/kabupaten K
P_P	: Jumlah penduduk di provinsi P
P_K	: Jumlah penduduk di kota/kabupaten K
\mathbf{y}	: Vektor variabel dependen berukuran $n \times 1$
\mathbf{X}	: Matriks berukuran $n \times (m + k + 1)$
$\boldsymbol{\beta}$: Vektor slope berukuran $k \times 1$
$\boldsymbol{\varepsilon}$: Vektor galat berukuran $n \times 1$
RLS	: Rata-rata lama sekolah
PPM	: Persentase penduduk miskin
f_i	: Penduduk 25 tahun ke atas untuk jenjang pendidikan ke- i
S_i	: Skor jenjang pendidikan ke- i
i	: Jenjang pendidikan, $i = 1, 2, 3, \dots, 21$