Arus Jurnal Sains dan Teknologi (AJST)
Website: http://jurnal.ardenjaya.com/index.php/ajst
Email: jurnal.ajst@ardenjaya.com

Estimasi Tingkat Pengangguran Terbuka di Pulau Sulawesi Menggunakan Regresi Nonparametrik Kernel dengan Optimasi *Bandwidth* SROT

INFO PENPULIS

Andi Tenri Ampa
Universitas Halu Oleo
anditenri.ampa@uho.ac.id

*Makkulau
Universitas Halu Oleo
*kulau.statistika@gmail.com

Adistya Meka Sahabaru
Universitas Halu Oleo
adistyamekasahabaru@gamil.com

INFO ARTIKEL

ISSN: 3026-3603
Vol. 4, No. 2 Oktober 2026
http://jurnal.ardenjaya.com/index.php/ajst

© 2026 Arden Jaya Publisher All rights reserved

Saran Penulisan Referensi:

Ampa, A. T., Makkulau, & Sahabaru, A. M. (2026). Estimasi Tingkat Pengangguran Terbuka di Pulau Sulawesi Menggunakan Regresi Nonparametrik Kernel dengan Optimasi Bandwidth SROT. *Arus Jurnal Sains dan Teknologi , 4* (2), 280-290.

Abstrak

Tingkat Pengangguran Terbuka (TPT) merupakan indikator utama dalam mengevaluasi kondisi ketenagakerjaan dan keberhasilan pembangunan ekonomi suatu wilayah. Secara agregat, TPT di Pulau Sulawesi pada tahun 2023 menunjukkan tren penurunan. Namun, disparitas yang cukup signifikan antar kabupaten/kota mengindikasikan adanya heterogenitas faktor-faktor yang memengaruhinya. Penelitian ini bertujuan untuk memodelkan TPT di Pulau Sulawesi menggunakan pendekatan regresi nonparametrik kernel Gaussian.Data yang digunakan merupakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS) tahun 2023 dengan unit observasi kabupaten/kota se-Pulau Sulawesi. Variabel respon dalam penelitian ini adalah Tingkat Pengangguran Terbuka (TPT), sedangkan variabel prediktor meliputi kepadatan penduduk, persentase penduduk miskin, pengeluaran per kapita, dan harapan lama sekolah.Metode analisis yang digunakan adalah regresi nonparametrik dengan estimator Nadaraya-Watson dan fungsi kernel Gaussian. Pemilihan metode ini didasarkan pada hasil uji linearitas yang menunjukkan adanya pola hubungan nonlinier antara variabel respon dan variabel prediktor. Untuk memperoleh estimasi yang optimal, penentuan parameter bandwidth dilakukan dengan menggunakan metode Silverman's Rule of Thumb (SROT).Hasil penelitian menunjukkan bahwa model regresi nonparametrik kernel Gaussian menghasilkan performa yang sangat baik dengan nilai koefisien determinasi (R2) sebesar 96% dan nilai Mean Squared Error (MSE) sebesar 0,15. Hal ini menunjukkan bahwa model mampu memberikan prediksi yang akurat serta menggambarkan pola hubungan antara TPT dan faktor-faktor yang memengaruhinya secara optimal.

Kata Kunci: Tingkat Pengangguran Terbuka; Regresi Nonparametrik; Kernel Gaussian; Nadaraya-Watson; Bandwidth SROT; Sulawesi

Abstract

Open Unemployment Rate (OUR) is a key indicator for evaluating labor market conditions and the success of regional economic development. In aggregate, the OUR in Sulawesi Island in 2023 showed a declining trend. However, significant disparities among regencies and municipalities indicate heterogeneity in the underlying determinant factors. This study aims to model the OUR in Sulawesi Island using a Gaussian kernel nonparametric regression approach.This study utilized secondary data sourced from Statistics Indonesia (BPS) in 2023, with regencies and municipalities in Sulawesi Island as the observation units. The response variable was the Open Unemployment Rate (OUR), while the predictor variables included population density, percentage of poor population, per capita expenditure, and expected years of schooling.The analytical method employed was nonparametric regression using the Nadaraya-Watson estimator with a Gaussian kernel function. This method was selected based on the results of a linearity test, which revealed a nonlinear relationship between the response variable and the predictor variables. To obtain an optimal estimation, the bandwidth parameter was determined using Silverman's Rule of Thumb (SROT) method.The results demonstrated that the Gaussian kernel nonparametric regression model performed excellently, with a coefficient of determination (R2) of 96% and a Mean Squared Error (MSE) of 0.15. This indicates that the model is capable of providing accurate predictions and optimally describing the relationship pattern between OUR and its influencing factors.

Keywords: Open Unemployment Rate; Nonparametric Regression; Gaussian Kernel; Nadaraya-Watson; SROT Bandwidth; Sulawesi.

A. Pendahuluan

Tingkat Pengangguran Terbuka (TPT) merupakan indikator yang digunakan untuk mengukur tingkat pengangguran. Pengangguran masih menjadi salah satu permasalahan ekonomi di Indonesia, termasuk di Pulau Sulawesi. TPT menunjukkan persentase jumlah pencari kerja terhadap total angkatan kerja. Tingginya angka pengangguran terbuka memiliki implikasi sosial yang luas karena mereka yang tidak bekerja tidak memiliki penghasilan. Semakin tinggi angka pengangguran terbuka, semakin besar potensi kerentanan sosial yang dapat menyebabkan, misalnya, kejahatan. Sebaliknya, semakin rendah angka pengangguran terbuka, semakin stabil kondisi sosial di masyarakat. Sangat tepat jika pemerintah sering menggunakan indikator ini sebagai ukuran keberhasilan pembangunan (Ampa *et al*., 2021).

Pada tahun 2023, TPT di beberapa provinsi di Pulau Sulawesi menunjukkan tren yang cenderung menurun dibandingkan tahun sebelumnya. TPT Provinsi Sulawesi Tenggara tercatat sebesar 3,66%, menurun sebesar 0,20% dibandingkan tahun 2022. Provinsi Sulawesi Barat sebesar 2,27% mengalami penurunan sebesar 0,07%, Sulawesi Tengah sebesar 2,95% turun 0,05%, Sulawesi Selatan sebesar 4,33% menurun 0,18% dan Sulawesi Utara sebesar 6,10% mengalami penurunan terbesar, yakni 0,51%. Berbeda dengan provinsi lainnya, Gorontalo mencatat TPT sebesar 3,06% atau meningkat sebesar 0,48% dibandingkan tahun 2022 (Badan Pusat Statistik, 2021a; 2021b; 2023ª; 2023b; & 2025).

Dalam regresi nonparametrik, khususnya untuk kasus multivariabel dimungkinkan untuk memiliki beberapa jenis estimator. Beberapa estimator dalam regresi nonparametrik, termasuk kernel digunakan untuk data tanpa pola, *Spline Smoothing* data berpola yang tidak bergantung pada titik simpul, *Fourier Series* dan *Spline Truncated*. Estimator kernel lebih sering digunakan dalam regresi nonparametrik. Hal ini karena estimator kernel lebih sederhana. Selain itu, estimator kernel memiliki kecepatan konvergensi yang relatif lebih cepat dibandingkan dengan estimator polinomial lokal, *deret fourier*, atau *spline* (Ampa *et al*., 2021; Anisa *et al*., 2019; Anwar, *et al*., 2025; Hardle, 1990; Laome, 2010; & Wulandary & Purnama, 2020).

Tujuan penelitian ini adalah mendapatkan model Tingkat Pengangguran Terbuka di Pulau Sulawesi menggunakan Regresi Nonparametrik Kernel *Gaussian*, dengan variabel Kepadatan Penduduk, Persentase Penduduk Miskin, Pengeluaran Per Kapita, dan Pengeluaran Per Kapita.

B. Metodologi

Penelitian ini dengan pendekatan kuantitatif yang lebih memfokuskan kemampuan dalam mengolah data serta angka-angka menggunakan statistik. Sumber data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari publikasi Badan Pusat Statistik

Provinsi Sulawesi Tenggara, Sulawesi Selatan, Sulawesi Barat, Sulawesi Tengah, Sulawesi Utara dan Gorontalo.

Variabel-variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

y : Tingkat Pengangguran Terbuka (Persen)

x_1 : Kepadatan Penduduk (Jiwa/km^2)

x_2 : Persentase Penduduk Miskin (Persen)

x_3 : Pengeluaran Per Kapita (Juta Rupiah)

x_4 : Pengeluaran Per Kapita (Tahun).

Pada penelitian ini menggunakan *software R-Studio*. Berikut langkah-langkah yang dilakukan pada penelitian ini.

1. Mengkaji literatur tentang TPT dan regresi nonparametrik kernel *Gaussian*.
2. Mengambil data mengenai TPT, Kepadatan Penduduk, Persentase Penduduk Miskin, Pengeluaran Per Kapita dan Harapan Lama Sekolah di *website* BPS tahun 2023.
3. Melakukan analisis statistik deskriptif.
4. Mengidentifikasi data dengan membuat diagram pencar (*scatterplot*) antara variabel respon dengan masing-masing variabel prediktor untuk mengetahui pola data.
5. Melakukan uji linearitas RESET *Test* antara variabel respon dengan masing-masing variabel prediktor yang bertujuan untuk mengetahui gambaran pola data masing-masing variabel.
6. Membentuk model regresi nonparametrik kernel *Gaussian* dengan estimator *Nadaraya-Watson*.
7. Menentukan *bandwidth* (h) menggunakan metode *Silverman's Rule of Thumb*
8. Melakukan pemodelan TPT menggunakan *bandwidth* optimal dengan fungsi regresi nonparametrik kernel *Gaussian* pada berbagai kombinasi variabel prediktor.
9. Menentukan model terbaik dari setiap kombinasi berdasarkan kriteria nilai *Mean Squared Error* (MSE) terkecil.
10. Interpretasi model terbaik.
11. Mengambil Kesimpulan.

C. Hasil dan Pembahasan

Statistik Deskriptif

Pada penelitian ini, data yang digunakan diperoleh dari *website* resmi Badan Pusat Statistik (BPS) Provinsi Sulawesi Tenggara, Sulawesi Selatan, Sulawesi Utara, Sulawesi Tengah, Sulawesi Barat dan Gorontalo tahun 2023. Penyajian data dengan metode statistik deskriptif menampilkan nilai rata-rata, simpangan baku, minimum, maksimum, kuartil 1, kuartil 3 dan interkuartil. Data yang diperoleh merupakan wilayah kab/kota yang ada di Pulau Sulawesi. Berikut adalah tabel statistik deskriptif (Chamdani, 2023 & Silverman, 1986) tiap variabel penelitian:

Tabel 1 Statistika Deskriptif

Var	Mean	SD	Min	Max	Q1	Q3	IQR
y	3,57	1,88	1,16	10,60	2,35	4,17	1,82
x_1	399,05	1.01 9,86	14,00	8227,00	55,00	283,00	228,00
x_2	10,91	3,78	4,59	18,38	7,42	14,03	6,61
x_3	10,70	2,04	7,56	17,89	9,04	11,95	2,91
x_4	13.23	0.94	11.81	16.91	12.59	13.50	0.91

Tabel 1 menyajikan karakteristik variabel respon Tingkat Pengangguran Terbuka (TPT) (y) di Pulau Sulawesi serta variabel prediktor yang meliputi Kepadatan Penduduk (x_1), Persentase Penduduk Miskin (x_2), Pengeluaran Per Kapita (x_3) dan Harapan Lama Sekolah (x_4). Berdasarkan Tabel 1 rata-rata TPT di Pulau Sulawesi sebesar 3,57% dengan simpangan baku 1,88. Nilai ini menunjukkan bahwa secara umum tingkat pengangguran berada pada kisaran tersebut dengan variasi yang relatif cukup besar antarwilayah, yaitu sekitar 1,88 poin persentase dari nilai rata-ratanya. TPT terendah sebesar 1,16% terdapat di Kabupaten Bombana, Provinsi Sulawesi Tenggara, sedangkan TPT tertinggi sebesar 10,6% berada di Kota Makassar, Provinsi Sulawesi Selatan. Tingginya TPT di Kota Makassar disebabkan oleh tingginya arus pencari kerja dari berbagai daerah yang tidak diimbangi dengan ketersediaan lapangan kerja yang memadai, serta karakteristik pasar kerja perkotaan yang kompetitif dan didominasi sektor formal. Sebaliknya, rendahnya TPT di Kabupaten Bombana dipengaruhi oleh dominasi sektor informal dan sektor primer seperti pertanian, perkebunan, dan pertambangan yang mampu menyerap tenaga kerja lokal, sehingga sebagian besar penduduk tetap dikategorikan bekerja meskipun produktivitasnya relatif rendah.

Variabel Kepadatan Penduduk (x_1) di Pulau Sulawesi memiliki rata-rata sebesar 399,05 jiwa/km^2 dengan simpangan baku 1.019,86. Nilai simpangan baku yang jauh lebih besar dari rata-rata menunjukkan adanya variasi yang sangat tinggi antar wilayah, sehingga distribusi kepadatan penduduk tidak merata. Hal ini mengindikasikan adanya ketimpangan yang signifikan antara daerah berpenduduk jarang dan daerah yang sangat padat, khususnya wilayah perkotaan. Kepadatan penduduk tertinggi terdapat di Kota Makassar, Provinsi Sulawesi Selatan, sebesar 8.227,00 jiwa/km^2, sedangkan yang terendah berada di Kabupaten Morowali Utara, Provinsi Sulawesi Tengah, sebesar 14,00 jiwa/km^2.

Variabel Persentase Penduduk Miskin (x_2) di Pulau Sulawesi memiliki rata-rata sebesar 10,91% dengan simpangan baku 3,78. Nilai tersebut menunjukkan bahwa secara umum sekitar 11 dari setiap 100 penduduk berada dalam kondisi miskin, dengan variasi antar wilayah yang cukup nyata, yakni sekitar 3–4 poin persentase dari rata-ratanya. Perbedaan ini mencerminkan ketimpangan tingkat kemiskinan di berbagai daerah. Persentase penduduk miskin tertinggi terdapat di Kabupaten Boalemo, Provinsi Gorontalo, sebesar 18,38%, sedangkan yang terendah berada di Kota Kendari, Provinsi Sulawesi Tenggara, sebesar 4,59%.

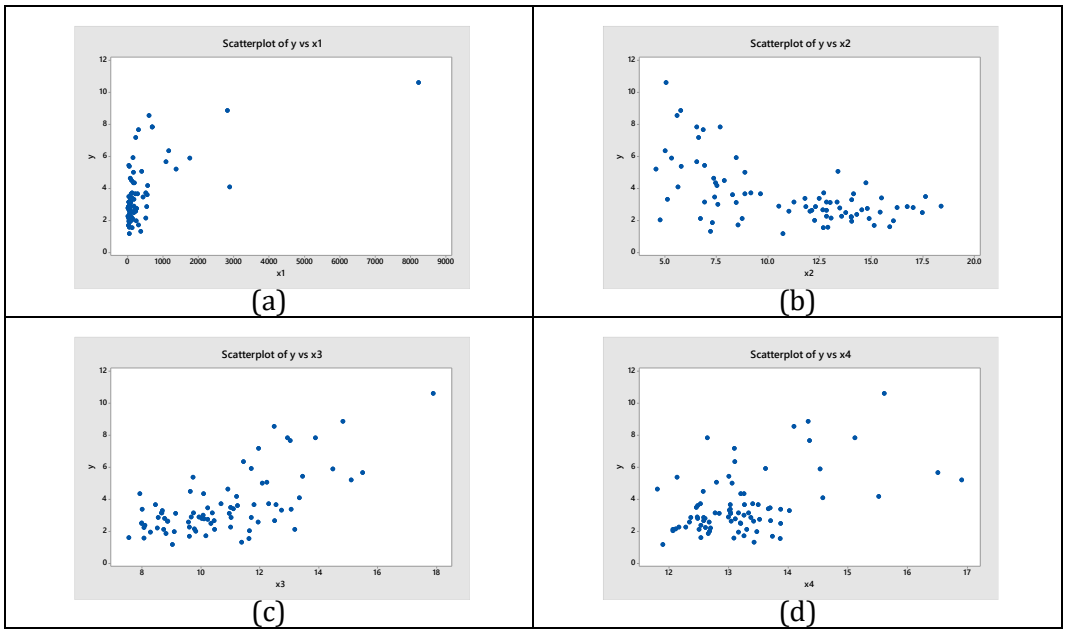
Variabel Pengeluaran Per Kapita (x_3) di Pulau Sulawesi memiliki rata-rata sebesar 10,70 artinya pengeluaran perkapita sebesar Rp.10.700.000 per bulan dengan simpangan baku 2,04 atau Rp.2.040.000. Perbedaan ini mencerminkan adanya ketimpangan tingkat kesejahteraan ekonomi masyarakat di berbagai daerah. Pengeluaran per kapita tertinggi terdapat di Kota Makassar, Provinsi Sulawesi Selatan, sebesar Rp.17.890.000, sedangkan yang terendah berada di Kabupaten Konawe Kepulauan, Provinsi Sulawesi Tenggara, sebesar Rp.7.560.000.

Variabel Harapan Lama Sekolah (x_4) di Pulau Sulawesi memiliki rata-rata sebesar 13,22 $\approx$ 13 tahun, yang menunjukkan bahwa secara umum penduduk diperkirakan menempuh pendidikan formal hingga sekitar jenjang Sekolah Menengah Atas (SMA). Simpangan baku sebesar 0,93 mengindikasikan bahwa variasi antarwilayah relatif kecil, sehingga kesempatan menempuh pendidikan formal tergolong cukup merata. Harapan Lama Sekolah tertinggi terdapat di Kota Kendari, Provinsi Sulawesi Tenggara, sebesar 16,91 $\approx$ 17 tahun, sedangkan yang

terendah berada di Kabupaten Bolaang Mongondow, Provinsi Sulawesi Utara, sebesar 11,81 ≈ 12 tahun.

Scatterplot

Sebelum melakukan analisis lebih lanjut terlebih dahulu diuji apakah variabel yang digunakan memenuhi asumsi linearitas dalam parameter atau tidak. Asumsi ini dapat menggunakan pendekatan grafis untuk mendeteksi apakah variabel prediktor yang digunakan memiliki hubungan yang linear atau tidak dengan variabel responnya. Berikut adalah bentuk dari *scatterplot* untuk keempat variabel prediktor dengan variabel respon.



Gambar 1 *Scatterplo*t Variabel Prediktor terhadap Variabel Respon

Secara visual Gambar 1 (a) terlihat bahwa bentuk pola hubungan antara variabel TPT (y) dan Kepadatan Penduduk (x_1) tidak membentuk pola tertentu, begitupun pada Gambar 1 (b), (c) dan (d) menunjukkan bahwa bentuk pola hubungan antara variabel TPT (y) dan variabel Persentase Penduduk Miskin (x_2), Pengeluaran Per Kapita (x_3), dan HLS (x_4) juga tidak mempunyai bentuk pola hubungan atau cenderung tidak mengikuti pola tertentu, sehingga dapat dimodelkan secara nonparametrik.

Uji Linearitas RESET *Test*

Uji linearitas akan memberikan penilaian yang lebih jelas mengenai kecocokan model regresi yang digunakan untuk mengestimasi hubungan antar variabel. Adapun uji linearitas yang digunakan adalah Uji *Ramsey* RESET *Test* untuk masing-masing variabel prediktor terhadap variabel respon:

Tabel 2 Uji Linearitas RESET Test

Variabel Respon	Variabel Prediktor	RESET *Test*	F Tabel	Keterangan
y	x_1	7,6887	3,12	Nonlinear
y	x_2	4,5128	3,12	Nonlinear
y	x_3	3,8163	3,12	Nonlinear
y	x_4	3,7757	3,12	Nonlinear

Berikut hipotesis uji linearitas:

H_0 : hubungan antara variabel prediktor dan variabel respon adalah linear.

H_1 : hubungan antara variabel prediktor dan variabel respon adalah nonlinear.

Pada Tabel 2, dapat dilihat bahwa nilai F_{hitung} pada variabel x_1, x_2, x_3, dan x_4 lebih besar dari F_{tabel} . Berdasarkan asumsi jika $F_{hitung} \geq F_{tabel}$, maka menolak H_0 , artinya model tidak memenuhi asumsi linearitas atau bersifat nonlinear, sehingga model menunjukkan bahwa hubungan antara *y* dan variabel-variabel tersebut bersifat nonlinear. Karena hubungan antara variabel respon dan variabel prediktor nonlinear dan tidak memiliki pola tertentu, maka estimasi yang digunakan adalah regresi nonparametrik.

Membentuk Model Regresi Nonparametrik Kernel *Gaussian* dengan Estimator *Nadaraya-Watson*

Berdasarkan analisis awal yang telah dilakukan diperoleh informasi bahwa hubungan antara variabel prediktor dan variabel respon tidak menunjukkan pola tertentu. Oleh karena itu, pendekatan regresi parametrik kurang sesuai digunakan, sehingga estimasi lebih tepat dilakukan dengan regresi nonparametrik. Karena variabel prediktor yang digunakan lebih dari 1 variabel maka digunakan regresi nonparametrik kernel *multivariable* diperoleh persamaan berikut.

$$y_i = m\left(x_{1i}, x_{2i}, .x_{3i}, x_{4i}\right) + \varepsilon_i \,;\; i = 1, 2, \ldots, 81 \qquad (1)$$

Data berupa $\left(x_{1i}, x_{2i}, x_{3i}, x_{4i}; y_i\right)$ untuk *i* =1, 2, 3, ..., 81.

$y_i = \sum_{i=1}^{n} m\left(x_{ji}\right) + \varepsilon_i$ untuk *i* = 1, 2, 3, ..., 81 dan *j* = 1, 2, 3, 4.

Selanjutnya fungsi $m\left(x_{ji}\right)$ dapat didekati dengan fungsi kernel, kemudian memasukan fungsi kernel kedalam model persamaan (4.1) sehingga diperoleh pesamaan berikut.

$$\hat{m}\left(x_{1i}, x_{2i}, x_{3i}, x_{4i}\right) = \sum_{j=1}^{4} \frac{\sum_{i=1}^{81} K_{hj}\left(\frac{x_j - x_{ji}}{h_j}\right) y_i^*}{\sum_{i=1}^{81} K_{hj}\left(\frac{x_j - x_{ji}}{h_j}\right)} \qquad (2)$$

Fungsi kernel *Gaussian* dipilih karena *range* pada fungsi ini berada dari $-\infty$ sampai ∞ dimana telah sesuai dengan data. Fungsi kernel *Gaussian* kemudian memasukkan fungsi *Gaussian* ke dalam fungsi kernel pada persamaan (2), sehingga diperoleh persamaan berikut:

$$\hat{m}\left(x_{1i}, x_{2i}, x_{3i}, x_{4i}\right) = \sum_{j=1}^{4} \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_j - x_{ji}}{h_j}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_j - x_{ji}}{h_j}\right)^2\right)} \qquad (3)$$

Sehingga diperoleh model sebagai berikut:

$$\hat{m}(x_{1i}, x_{2i}, x_{3i}, x_{4i}) = \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{h_1}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{h_1}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{h_2}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{h_2}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_3 - x_{3i}}{h_3}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_3 - x_{3i}}{h_3}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_4 - x_{4i}}{h_4}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_4 - x_{4i}}{h_4}\right)^2\right)} \quad (4)$$

Penentuan *Bandwidth* Optimal

Hal yang paling penting dalam memperoleh estimasi kurva regresi dengan pendekatan kernel adalah pemilihan *bandwidth* yang tepat, karena kualitas estimator kernel sangat bergantung pada nilai *bandwidth* dan berpengaruh terhadap model yang dihasilkan. Nilai *bandwidth* yang terlalu besar menghasilkan *plot* yang sangat halus (*over smooth*), jika nilai *bandwidth* yang digunakan terlalu kecil, maka akan menghasilkan *plot* yang sangat kasar (*under smooth*), sehingga perlu ditentukan *bandwidth* paling optimal. Adapun pada penelitian ini penentuan *bandwidth* optimal menggunakan Aturan Silverman diperoleh (Silverman, 1986):

$$h = 0,9 \times \min\left\{S, \frac{IQR}{1,34}\right\} \times n^{-\frac{1}{5}}$$

Untuk menghitung *bandwidth* variabel pertama (h_1):

$$h_1 = 0,9 \times \min\left\{1.019,86, \frac{228,00}{1,34}\right\} \times 81^{-\frac{1}{5}}$$

$$h_1 = 0,9 \times \min\left\{1.019,86,\ 170,14\right\} \times 81^{-\frac{1}{5}}$$

$$h_1 = 0,9 \times 170,14 \times 0,415$$

$$h_1 = 63,59$$

dengan menggunakan rumus dan langkah yang sama di hitung *bandwidth* untuk setiap variabel, diperoleh h_2 = 1,41, h_3 = 0,76, dan h_4 = 0,25.

Pemodelan Tingkat Pengangguran Terbuka (TPT) di Pulau Sulawesi menggunakan Regresi Nonparametrik Kernel *Gaussian* dengan Beberapa Kombinasi Variabel

Estimasi parameter yang telah diperoleh dengan komponen nonparametrik menggunakan metode *bandwidth Silverman Rhule of Thumb* (SROT) kemudian dimasukkan kedalam model. Untuk memodelkan Tingkat Pengangguran Terbuka (TPT) di Pulau Sulawesi, dilakukan analisis dengan menggunakan beberapa kombinasi variabel prediktor terhadap variabel respon, nilai *Mean Squared Error* (MSE) dan koefisien determinasi $\left(R^2\right)$. Model terbaik dilihat dengan nilai MSE terkecil, secara lengkap disajikan pada Tabel 3.

Tabel 3 Model Estimasi dari Beberapa Kombinasi Variabel Prediktor terhadap Variabel Respon dengan MSE dan R^2

Kombinasi	Variabel	*Mean Squared Error* (MSE)	R^2
4 Variabel	x_1, x_2, x_3, x_4	**0,15**	**0,96**
3 Variabel	x_1, x_2, x_3	**0,33**	**0,90**
	x_1, x_2, x_4	0,49	0,86
	x_1, x_3, x_4	0,51	0,85
	x_2, x_3, x_4	0,52	0,85
2 Variabel	x_1, x_2	**0,86**	**0,76**
	x_1, x_3	0,87	0,75
	x_1, x_4	0,89	0,75
	x_2, x_3	1,32	0,62
	x_2, x_4	1,08	0,69
	x_3, x_4	1,12	0,68
1 Variabel	x_1	**1,47**	**0,58**
	x_2	2,15	0,39
	x_3	1,71	0,51
	x_4	2,06	0,41

Pada Tabel 3 model beberapa kombinasi variabel tersebut dapat ditulsikan sebagai berikut:

1. Model dengan 4 Kombinasi Variabel (x_1, x_2, x_3, x_4)

$$\hat{H}(x_{1i}, x_{2i}, x_{3i}, x_{4i}) = \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{1,41}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{1,41}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_3 - x_{3i}}{0,76}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_3 - x_{3i}}{0,76}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_4 - x_{4i}}{0,25}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_4 - x_{4i}}{0,25}\right)^2\right)} \quad (5)$$

2. Model dengan 3 Kombinasi Variabel (x_1, x_2, x_3)

$$\hat{H}(x_{1i}, x_{2i}, x_{3i}) = \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{1,41}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{1,41}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_3 - x_{3i}}{0,76}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_3 - x_{3i}}{0,76}\right)^2\right)} \quad (6)$$

3. Model dengan 2 Kombinasi Variabel (x_1, x_2)

$$\hat{H}(x_{1i}, x_{2i}) = \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{1,41}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{1,41}\right)^2\right)} \quad (7)$$

4. Model dengan 1 Kombinasi Variabel (x_1)

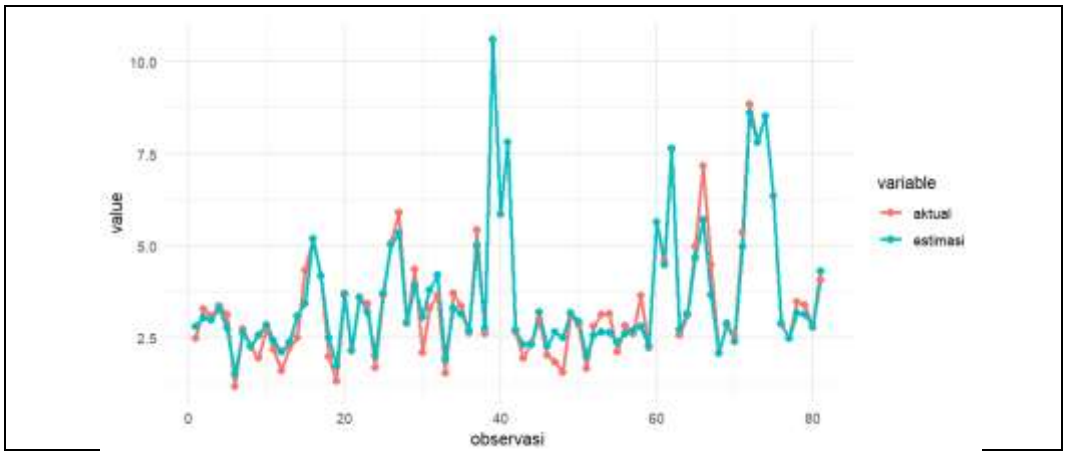
$$\hat{H}(x_{1i}) = \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right) y_{i^*}}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right)} \quad (8)$$

Berdasarkan Tabel 3 dan uraian model, kombinasi 4 variabel (x_1, x_2, x_3, x_4) menghasilkan nilai MSE terkecil sebesar 0,15 dan koefisien determinasi (R^2) sebesar 0,96. Hal ini menunjukkan bahwa model memiliki tingkat kesalahan prediksi yang rendah serta mampu menjelaskan sebesar 96% variasi pada variabel respon. Dengan demikian, model dengan 4 variabel merupakan model terbaik karena memiliki kemampuan prediksi yang sangat baik dibandingkan dengan kombinasi variabel lainnya.

Berdasarkan hasil pemilihan model, diperoleh bahwa model terbaik adalah model dengan menggunakan 4 variabel prediktor. Oleh karena itu, dalam penelitian ini digunakan model regresi nonparametrik dengan estimator kernel *Gaussian* serta penentuan *bandwidth* optimal untuk memodelkan TPT di Pulau Sulawesi. Model yang digunakan dinyatakan dalam bentuk persamaan 5. Dari model yang dihasilkan pada persamaan 5, sehingga nilai *error* dapat hitung. Secara lengkap dapat dilihat pada Tabel 4.

Tabel 4 Perbandingan *y* Aktual dan *y* Estimasi

No.	Kabupaten/Kota	*y* Aktual	*y* Estimasi	*Error*
1.	Buton	2,48	2,80	-0,32
2.	Muna	3,27	3,03	0,24
3.	Konawe	3,09	2,98	0,11
4.	Kolaka	3,36	3,30	0,06
5.	Konawe Selatan	3,12	2,75	0,37
6.	Bombana	1,16	1,48	-0,32
7.	Wakatobi	2,73	2,63	0,10
8.	Kolaka Utara	2,24	2,25	-0,01
9.	Buton Utara	1,93	2,55	-0,62
10.	Konawe Utara	2,76	2,84	-0,08
11.	Kolaka Timur	2,18	2,42	-0,24
12.	Konawe Kepulauan	1,59	2,12	-0,53
13.	Muna Barat	2,22	2,37	-0,15
...				
77.	Gorontalo	2,48	2,48	0
78.	Pohuwato	3,46	3,16	0.3
79.	Bone Bolango	3,39	3,12	0,27
80.	Gorontalo Utara	2,77	2,80	-0,03



Gambar 2 Grafik Perbandingan y Aktual vs y Estimasi

Berdasarkan Gambar 2, secara visual terlihat bahwa data respon aktual dengan data respon prediksi memiliki pola data yang relatif sama, artinya data prediksi mampu mengikuti data aktual. Meskipun terdapat beberapa penyimpangan pada observasi tertentu, secara umum selisih antara nilai aktual dan estimasi relatif kecil, sehingga dapat disimpulkan bahwa model memiliki kemampuan prediksi yang baik. Hal tersebut dapat dinyatakan bahwa pemodelan yang diperoleh telah sesuai untuk memprediksi Tingkat Pengangguran Terbuka di Pulau Sulawesi. Wilayah dengan nilai prediksi tertinggi yaitu berada di Kota Makassar, Provinsi Sulawesi Selatan sedangkan wilayah dengan nilai prediksi terendah berada di Kabupaten Bombana, Provinsi Sulawesi Tenggara.

D. Kesimpulan

Berdasarkan hasil analisis pemodelan tingkat pengangguran terbuka di Pulau Sulawesi menggunakan regresi nonparametrik kernel *Gaussian* dengan menggunakan estimator *Nadaraya-Watson* maka didapatkan model terbaik yaitu:

$$\hat{m}(x_{1i}, x_{2i}, x_{3i}, x_{4i}) = \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right) y_i^*}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_1 - x_{1i}}{63,59}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{1,41}\right)^2\right) y_i^*}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_2 - x_{2i}}{1,41}\right)^2\right)} +$$
$$+ \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_3 - x_{3i}}{0,76}\right)^2\right) y_i^*}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_3 - x_{3i}}{0,76}\right)^2\right)} + \frac{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_4 - x_{4i}}{0,25}\right)^2\right) y_i^*}{\sum_{i=1}^{81} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x_4 - x_{4i}}{0,25}\right)^2\right)}$$

Pemodelan regresi nonparametrik kernel *Gaussian* diperoleh nilai *Mean Squared Error* (MSE) sebesar 0,15. Nilai tersebut menunjukkan bahwa model memiliki tingkat kesalahan prediksi yang relatif kecil, sehingga dapat dikatakan bahwa model memiliki kemampuan prediksi yang sangat baik. Selain itu, nilai koefisien determinasi (R^2) sebesar 96% menunjukkan bahwa 96% variasi pada variabel respon dapat dijelaskan oleh variabel prediktor dalam model.

E. Referensi

Ampa, T.A., Budiantara, I.N., & Zain, I. (2021). Kernel Multivariable Semiparametric Regression Model in Estimating the Level of Open Unemption East Java Province. *Journal of Physics: Conference Series, 1899* (012127).

Anisa, N., Debataraja, N.N., & Martha, S. (2019). Estimasi Model Regresi Nonparametrik Kernel Menggunakan Estimator Nadaraya-Watson. *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster), 08*(4), 633-638.

Anwar, R., Rassiyanti, L., & Pitri, R. (2025). Analisis Regresi Kernel Gaussian untuk Memprediksi Indeks Pembangunan Manusia (IPM) Berdasarkan Faktor Sosial-Ekonomi Provinsi di Indonesia. *Jurnal Riset Rumpun Matematika dan Ilmu Pengetahuan Alam, 4*(3), 44-53.

Badan Pusat Statistik Provinsi Sulawesi Barat (BPS). (2021a). *Statistika Angkatan Kerja Provinsi Sulawesi Barat 2021.* Kab. Mamuju: Badan Pusat Statistik Provinsi Sulawesi Barat.

Badan Pusat Statistik (BPS). (2021b). *Booklet SAKERNAS (Survei Angkatan Kerja Nasional).* Jakarta Pusat: Badan Pusat Statistik.

Badan Pusat Statistik (BPS). (2023a). Tingkat Pengangguran Terbuka (TPT).

Badan Pusat Statistik (BPS). (2023b). *Keadaan Angkatan Kerja di Indonesia. Jakarta Pusat*: Badan Pusat Statistik.

Badan Pusat Statistik (BPS). (2025). *Statistik Indonesia 2025.* Jakarta Pusat: Badan Pusat Statistik.

Chamdani, M.N. (2023). Statistika Deskriptif: Bilangan Rangkuman. *Jurnal Statistika*.

Hardle, W. (1990). *Applied Nonparametric Regression*. New York: Cambridge University Press.

Laome, L. (2010). Perbandingan Model Regresi Nonparametrik dengan Spline dan Kernel. *JIMT, 7*(1).

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. New York: Monographs on Statistics and Applied Probability.

Wulandary, S. & Purnama, D.I. (2020). Perbandingan Regresi Nonparametrik Kernel NWE dan B-Splines pada Pemodelan Rata-rata Lama Sekolah dan Pengeluaran Perkapita di Indonesia. *Jambura Journal of Probability and Statistics, 1*(2).