

**PEMODELAN PERSENTASE PENDUDUK MISKIN DI JAWA  
TIMUR MENGGUNAKAN REGRESI NONPARAMETRIK  
DENGAN PENDEKATAN KERNEL GAUSSIAN**

**Skripsi**

Disusun sebagai salah satu syarat  
untuk memperoleh gelar Sarjana Statistika

Program Studi Statistika



Oleh

Nailil Muna

2520200070

**UNUGIRI**

**PROGRAM STUDI STATISTIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS NAHDLATUL ULAMA SUNAN GIRI**

**BOJONEGORO**

**2023**

## 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.

Bojonegoro, 2 juli 2024



Nailil Muna  
NIM. 2520200070



# UNUGIRI

## HALAMAN PERSETUJUAN

Nama : Nailil Muna

Nim : 2520200070

Judul : Pemodelan Persentase Penduduk Miskin di Jawa Timur Menggunakan  
Regresi Nonparametrik dengan Pendekatan *Kernel Gaussian*

Telah disetujui dan dinyatakan memenuhi syarat untuk diajukan dalam ujian  
skripsi. Bojonegoro, 2 Juli 2024



# UNUGIRI

## HALAMAN PENGESAHAN

Nama : Nailil Muna  
Nim : 2520200070  
Judul : Pemodelan Persentase Penduduk Miskin di Jawa Timur Menggunakan Regresi Nonparametrik dengan Pendekatan Kernel Gaussian

Telah dipertahankan di hadapan penguji pada tanggal 10 Juli 2024

Dewan Penguji

Tim Pembimbing

Penguji I


Pembimbing I

  
Nur Mahmudah, M.Stat.  
NIDN. 0715039201

  
Alif Yuanita Kartini, M.Si.  
NIDN. 0721048606

Penguji II


Pembimbing II

  
Dr. H. M. Ridlwan Hambali, Lc., M. A.  
NIDN. 2117056803

  
Jauhara Raha Budiani, M.Stat  
NIDN. 0705079201


Mengetahui,

Dean Fakultas Sains dan Teknologi

  
M. Jauhar Vikri, M.Kom.  
NIDN. 0712078803

Mengetahui,

Ketua Program Studi

  
Nur Mahmudah, M.Stat.  
NIDN. 0715039201

STAFIKA  
FST UNUGIRI  
BESUKONGKO

## MOTTO

Tidak ada ujian yang tidak bisa diselesaikan. Tidak ada kesulitan yang melebihi batas kesanggupan. Karena Allah tidak akan membebani seseorang melainkan sesuai dengan kadar kesanggupannya.

(QS. Al- Baqarah : 286)

## PERSEMBAHAN

Bapak dan Ibuku tercinta

**Mochammad Maskuri (Bapak) dan Siti Fatimah (Ibu)**

*Terima kasih yang tiada terhingga untuk seluruh do'a dan cinta serta semua hal yang telah diberikan. Terima kasih telah menjadi orang tua terbaik dalam hidupku dan selalu menyisipkan namaku dalam setiap sujud*

Kakak dan Adik-adik tersayang

**Miftahul Khoir (kakak), Fatihatul Rizka (Adik), dan Ahmad Silakhudin (Adik)**

*Tiada yang paling mengharukan saat berkumpul bersama kalian, walaupun sering bertengkar tetapi hal ini akan selalu menjadi warna yang tidak akan bisa tergantikan. Terima kasih atas do'a dan bantuan kalian selama ini*

# UNUGIRI

## KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas berkah rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan proposal skripsi ini. terselesaikannya skripsi ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu penulis ingin menyampaikan terima kasih kepada:

1. Bapak K.M. Jauharul Ma'arif, M.d.I selaku Rektor Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
2. Bapak Muhammad Jauhar Vikri, M. Kom. selaku Dekan Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
3. Ibu Nur Mahmudah, M.Stat. selaku Ketua Program Studi Statistika dan Sains Data Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
4. Ibu Alif Yuanita, M.Si. selaku Dosen Pembimbing I penyusunan skripsi atas segala arahan dan bimbingan yang telah diberikan sehingga penyusunan skripsi ini dapat selesai dengan baik dan lancar.
5. Ibu Jauhara Rana Budiani, M.Stat. selaku Dosen Pembimbing II penyusunan skripsi yang sudah memberikan waktu bimbingan dan arahan selama penyusunan skripsi ini.
6. Bapak Denny Nurdiansyah, M.Si. selaku Dosen Pembimbing Akademik yang telah membimbing penulis dalam menyelesaikan studi.
7. Semua pihak yang sudah membantu penulis dalam menyelesaikan proposal skripsi sehingga tidak dapat disebutkan namanya satu persatu.

Penulis menyadari dalam penulisan skripsi ini tidak luput dari berbagai kekurangan. Penulis mengharapkan saran dan kritik demi kesempurnaan dan perbaikannya sehingga akhirnya skripsi ini dapat memberikan manfaat bagi bidang pendidikan dan penerapan di lapangan serta bisa dikembangkan lagi lebih lanjut.

Bojonegoro, 2 Juli 2024

Nailil Muna

## ***ABSTRACT***

Nailil Muna. 2024. Modeling the Percentage of Poor Population in East Java Using Nonparametric Regression with a Gaussian Kernel Approach. Thesis, Department of Statistics, Faculty of Science and Technology, Universitas Nahdlatul Ulama Sunan Giri Bojonegoro. Supervisor (I) Alif Yuanita Kartini, M.Si. (II) Jauhara Rana Budiani, M.Stat. Poverty is a major problem that various countries around the world want to solve. This is also the case in Indonesia, which has a poverty graph that is directly proportional to that of East Java. However, poverty in East Java still exceeds the poverty rate in Indonesia. So it is necessary to conduct a study to analyze the percentage of poor people in East Java in the hope that it can become a reference for the relevant government in overcoming poverty in East Java. This study uses nonparametric kernel regression with the Nadaraya-Watson estimator for data on the percentage of poor people in East Java. The kernel function used in this study is the Gaussian kernel function. The data used is sourced from the National Socio-Economic Survey (Susenas) Consumption and Expenditure Module data published by the Central Bureau of Statistics (BPS) of East Java Province. Includes data on the percentage of poor people and factors that affect the percentage of poor people, namely the open unemployment rate (TPT), life expectancy rate (AHH), and labor force participation rate (TPAK) in 2023. Based on the results of the study, it shows that the kernel nonparametric regression model with the Nadaraya-Watson estimator and Gaussian function produces an optimal Generalized Cross Validation (GCV) of 0.139558 with a bandwidth value of  $h_1=h_2=h_3=5$ . The accuracy of the best model obtained from the MAE (Mean Absolute Error) value of 0.187 which indicates that all predictor models obtained have good accuracy because the value is close to 0.

Keywords: Percentage of poor population, TPT, AHH, TPAK, Kernel nonparametric regression, Nadaraya-Watson, Gaussian Function Approach, Bandwidth



UNUGIRI

## ABSTRAK

Nailil Muna. 2024. Pemodelan Persentase Penduduk Miskin Di Jawa Timur Menggunakan Regresi Nonparametrik Dengan Pendekatan Kernel Gaussian. Skripsi, Jurusan Statistika Fakultas Sains dan teknologi, Universitas Nahdlatul Ulama Sunan Giri Bojonegoro. Pembimbing (I) Alif Yuanita Kartini, M.Si. (II) Jauhara Rana Budiani, M.Stat. Kemiskinan menjadi masalah utama yang ingin dituntaskan oleh berbagai negara di seluruh dunia. Begitu juga di Negara Indonesia yang memiliki grafik kemiskinan yang berbanding lurus dengan Jawa Timur. Namun kemiskinan di Jawa Timur masih melebihi angka kemiskinan di Indonesia. Sehingga perlu dilakukan suatu penelitian untuk menganalisis persentase penduduk miskin di Jawa Timur dengan harapan bisa menjadi suatu acuan untuk pemerintah terkait dalam mengatasi kemiskinan di Jawa Timur. Penelitian ini, menggunakan regresi nonparametri *kernel* dengan estimator *Nadaraya-Watson* untuk data Persentase penduduk miskin di Jawa Timur. Fungsi *kernel* yang digunakan dalam penelitian ini adalah fungsi *kernel Gaussian*. Data yang digunakan bersumber dari data Survei Sosial Ekonomi Nasional (Susenas) Modul Konsumsi dan Pengeluaran yang dipublikasikan oleh Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Meliputi data Persentase penduduk miskin dan faktor yang mempengaruhi Persentase penduduk miskin yaitu Tingkat pengangguran terbuka (TPT), Angka Harapan Hidup (AHH), serta Tingkat partisipasi angkatan kerja (TPAK) pada tahun 2023. Berdasarkan hasil penelitian menunjukkan bahwa model regresi nonparametrik *kernel* dengan estimator *Nadaraya-Watson* dan fungsi *Gaussian* menghasilkan *Generalized Cross Validation (GCV)* optimal sebesar 0.139558 dengan nilai *bandwidth*  $h_1=h_2=h_3=5$ . Keakuratan model terbaik yang diperoleh dari nilai (*Mean Absolute Error*) MAE sebesar 0,187 yang menunjukkan bahwa semua model prediktor yang diperoleh memiliki akurasi yang baik karena nilainya mendekati 0.

Kata Kunci : Persentase penduduk miskin, TPT, AHH, TPAK, Regresi nonparametrik *Kernel*, *Nadaraya-Watson*, Pendekatan Fungsi *Gaussian*, *Bandwidth*

UNUGIRI



# DAFTAR ISI

	Halaman
<b>HALAMAN SAMPUL LUAR (COVER)</b> .....	<b>i</b>
<b>HALAMAN SAMPUL DALAM</b> .....	<b>ii</b>
<b>HALAMAN PERNYATAAN</b> .....	<b>iii</b>
<b>HALAMAN PERSETUJUAN</b> .....	<b>iv</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>v</b>
<b>MOTTO DAN PERSEMBAHAN</b> .....	<b>vi</b>
<b>KATA PENGANTAR</b> .....	<b>vii</b>
<b>ABSTRACT</b> .....	<b>viii</b>
<b>ABSTRAK</b> .....	<b>ix</b>
<b>DAFTAR ISI</b> .....	<b>x</b>
<b>DAFTAR TABEL</b> .....	<b>xiii</b>
<b>DAFTAR BAGAN</b> .....	<b>xiv</b>
<b>DAFTAR GAMBAR</b> .....	<b>xv</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xvi</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1    Latar belakang .....	<b>1</b>
1.2    Rumusan Masalah.....	<b>5</b>
1.3    Tujuan Penelitian.....	<b>5</b>
1.4    Manfaat Penelitian .....	<b>5</b>
1.5    Batasan Permasalahan .....	<b>6</b>
<b>BAB II TINJAUAN PUSTAKA DAN DASAR TEORI</b> .....	<b>7</b>
2.1    Tinjauan Pustaka.....	<b>7</b>
2.2    Dasar Teori .....	<b>13</b>

2.2.1	Statistik Deskriptif .....	13
2.2.1.1	<i>Ukuran Pemusatan dan Penyebaran</i> .....	13
2.2.1.2	<i>Diagram dan Grafik</i> .....	14
2.2.1.3	<i>Pemetaan</i> .....	15
2.2.2	<i>Rescalling Data dan Scatter Plot</i> .....	16
2.2.2.1	<i>Rescaling Data</i> .....	16
2.2.2.2	<i>Scatter Plot</i> .....	17
2.2.3	Regresi Nonparametrik .....	17
2.2.3.1	<i>Regresi Nonparametrik Kernel</i> .....	19
2.2.3.2	<i>Estimasi Kernel Nadaraya-Watson</i> .....	20
2.2.3.3	<i>Fungsi Kernel Gaussian</i> .....	23
2.2.3.4	<i>Pemilihan Bandwidth Optimum</i> .....	25
2.2.4	Prediksi Model Regresi Nonparametrik Kernel Dengan Pendekatan Gaussian .....	25
2.2.4.1	<i>Uji Keakuratan Model</i> .....	26
2.2.4.2	<i>Pemetaan persentase penduduk miskin di Jawa Timur berdasarkan nilai prediksi</i> .....	26
2.2.5	Persentase kemiskinan .....	26
2.2.5.1	<i>Tingkat Pengangguran Terbuka (TPT)</i> .....	28
2.2.5.2	<i>Angka Harapan Hidup</i> .....	29
2.2.5.3	<i>Tingkat Partisipasi Angkatan Kerja</i> .....	29

**BAB III METODE PENELITIAN .....** 30

3.1	Sumber Data .....	30
3.2	Variabel Penelitian .....	30
3.3	Langkah-langkah Analisis .....	32
3.4	Diagram Alir .....	33

<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>35</b>
4.1 Statistika Deskriptif .....	35
4.2 Pemodelan Regresi Nonparametrik .....	43
4.2.1 Pola Hubungan Antara Variabel Dependen dengan Variabel Independen.....	43
4.2.2 Pemodelan Regresi Nonparametrik dengan Pendekatan <i>Kernel Gaussian</i> .....	47
4.3 Prediksi Persentase Penduduk Miskin berdasarkan <i>Bandwidth</i> Optimum.....	50
4.3.1 Hasil Prediksi Persentase Penduduk Miskin berdasarkan <i>Bandwidth</i> Optimum.....	50
4.3.2 Evaluasi Keakuratan Model Terbaik <i>Regresi Nonparametrik Kernel</i> Fungsi <i>Gaussian</i> .....	52
4.3.3 Hasil Pemetaan .....	54
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>57</b>
5.1 Kesimpulan .....	57
5.2 Saran .....	57
<b>DAFTAR PUSTAKA .....</b>	<b>58</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>62</b>

**UNUGIRI**

## DAFTAR TABEL

<b>Tabel 2.1</b> Penelitian terdahulu .....	7
<b>Tabel 2.2</b> Fungsi Kernel .....	23
<b>Tabel 3.1</b> Variabel Penelitian.....	31
<b>Tabel 3.2</b> Struktur Data Penelitian .....	31
<b>Tabel 4.1</b> Statistik Deskriptif Persentase penduduk miskin.....	35
<b>Tabel 4.2</b> Statistik Deskriptif Tingkat Pengangguran Terbuka .....	37
<b>Tabel 4.3</b> Statistik Deskriptif Angka harapan hidup .....	39
<b>Tabel 4.4</b> Statistik Deskriptif Tingkat Partisipasi Angkatan Kerja .....	41
<b>Tabel 4.5</b> Nilai Minimum dan Nilai Maksimum.....	44
<b>Tabel 4.6</b> Bandwidth $h_1$ , $h_2$ , dan $h_3$ dengan GCV.....	48
<b>Tabel 4.7</b> Nilai estimasi untuk $n = 38$ pada $h_1=5$ , $h_2=5$ , dan $h_3=5$ .....	50
<b>Tabel 4.8</b> Perbandingan nilai aktual hasil rescaling dengan nilai prediksi.....	52
<b>Tabel 4.9</b> Pengelompokan tingkat persentase penduduk miskin di Provinsi Jawa Timur berdasarkan model Kernel Gaussian .....	56

UNUGIRI

## DAFTAR BAGAN

**Bagan 3.1** Diagram Alir penelitian..... 34



## DAFTAR GAMBAR

<b>Gambar 4.1</b> Grafik Persentase penduduk miskin di Jawa Timur tahun 2023 .....	36
<b>Gambar 4.2</b> Pengelompokan Persentase penduduk miskin di Jawa Timur tahun 2023 .....	36
<b>Gambar 4.3</b> Grafik Tingkat Pengangguran Terbuka di Jawa Timur tahun 2023	38
<b>Gambar 4.4</b> Pengelompokan Tingkat Pengangguran Terbuka di Jawa Timur tahun 2023 .....	38
<b>Gambar 4.5</b> Grafik Angka harapan hidup di Jawa Timur tahun 2023 .....	40
<b>Gambar 4.6</b> Pengelompokan Angka harapan hidup di Jawa Timur tahun 2023 .....	40
<b>Gambar 4.7</b> Grafik Tingkat Partisipasi Angkatan Kerja di Jawa Timur tahun 2023 .....	42
<b>Gambar 4.8</b> Pengelompokan Tingkat Partisipasi Angkatan Kerja di Jawa Timur tahun 2023 .....	43
<b>Gambar 4.9</b> Plot Hubungan Variabel TPT ( $x_1$ ) dengan Variabel Persentase penduduk miskin ( $y$ ).....	45
<b>Gambar 4.10</b> Plot Hubungan Variabel AHH ( $x_2$ ) dengan Variabel Persentase penduduk miskin ( $y$ ).....	46
<b>Gambar 4.11</b> Plot Hubungan Variabel TPAK ( $x_3$ ) dengan Variabel Persentase penduduk miskin ( $y$ ).....	47
<b>Gambar 4.12</b> Perbandingan data aktual dengan data prediksi.....	54
<b>Gambar 4.13</b> Peta nilai prediksi data persentase penduduk miskin Daerah Jawa Timur .....	55

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Data Persentase Penduduk Miskin (Y), Tingkat Pengangguran Terbuka (X1), Angka Harapan Hidup (X3), dan Tingkat Partisipasi Angkatan Kerja.....	62
<b>Lampiran 2.</b> Hasil <i>Rescaling</i> Data Persentase Penduduk Miskin (Y), Tingkat Pengangguran Terbuka (X1), Angka Harapan Hidup (X3), dan Tingkat Partisipasi Angkatan Kerja (X3).....	64
<b>Lampiran 3.</b> Hasil <i>Bandwidth</i> dan <i>GCV</i> .....	66
<b>Lampiran 4.</b> Source Code Program Matlab.....	67



# UNUGIRI