

PENERAPAN MODEL *GEOGRAPHICALLY WEIGHTED LOGISTIC REGRESSION* (GWL_R) DENGAN FUNGSI PEMBOBOT *ADAPTIVE GAUSSIAN KERNEL* PADA DATA KEMISKINAN DI INDONESIA

(Skripsi)

Oleh

**TRI HASTUTI
NPM 1617031043**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022**

**PENERAPAN MODEL GEOGRAPHICALLY WEIGHTED LOGISTIC
REGRESSION (GWLR) DENGAN FUNGSI PEMBOBOT ADAPTIVE
GAUSSIAN KERNEL PADA DATA KEMISKINAN DI INDONESIA**

Oleh

TRI HASTUTI

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam



FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022

ABSTRAK

PENERAPAN MODEL GEOGRAPHICALLY WEIGHTED LOGISTIC REGRESSION (GWLR) DENGAN FUNGSI PEMBOBOT ADAPTIVE GAUSSIAN KERNEL PADA DATA KEMISKINAN DI INDONESIA

Oleh

TRI HASTUTI

Analisis regresi logistik merupakan analisis yang digunakan untuk mengetahui hubungan antara variabel respon yang bersifat kategori dengan satu atau lebih variabel prediktor dengan asumsi bahwa respon tidak dipengaruhi lokasi geografis (data spasial). Metode GWLR adalah bentuk lokal dari regresi logistik dimana lokasi geografis diperhatikan. Tujuan dari penelitian ini adalah untuk menerapkan model GWLR dengan fungsi pembobot *Adaptive Gaussian Kernel* guna mengetahui faktor-faktor yang mempengaruhi persentase kemiskinan di Indonesia menggunakan *Geographically Weighted Logistic Regression* (GWLR) dengan fungsi pembobot *Adaptive Gaussian Kernel*. Hasil dari penelitian ini didapatkan model GWLR dengan fungsi pembobot *Adaptive Gaussian Kernel* tiap Provinsi di Indonesia dan variabel Produk Domestik Regional Bruto (PDRB) per kapita (X_3) berpengaruh secara signifikan terhadap persentase kemiskinan di Indonesia tahun 2021.

Kata Kunci: Regresi Logistik, GWLR, dan *Adaptive Gaussian Kernel*

ABSTRACT

PENERAPAN MODEL GEOGRAPHICALLY WEIGHTED LOGISTIC REGRESSION (GWLR) DENGAN FUNGSI PEMBOBOT ADAPTIVE GAUSSIAN KERNEL PADA DATA KEMISKINAN DI INDONESIA

By

TRI HASTUTI

Logistic regression analysis is an analysis used to determine the relationship between categorical response variables and one or more predictor variables with the assumption that responses are not influenced by geographic location (spatial data). *Geographically Weighted Logistic Regression (GWLR)* is a local form of the logistic regression where geographical factors considered. The purpose of this study is to apply the GWLR model with *Adaptive Gaussian Kernel* weighted function to determine which factors affect the percentage of poverty in Indonesia, using *Geographically Weighted Logistic Regression (GWLR)* with *Adaptive Gaussian Kernel* weighted function. The results of this study obtained the GWLR model with *Adaptive Gaussian Kernel* weighted function for each province in Indonesia and variable Gross Regional Domestic Product (GRDP) per capita (X_3) has significantly affect the percentage of poverty in Indonesia 2021.

Keywords: Logistic Regression, GWLR, and *Adaptive Gaussian Kernel*

Judul Skripsi : Penerapan Model *Geographically Weighted Logistic Regression* (GWLR) dengan Fungsi Pembobot *Adaptive Gaussian Kernel* Pada Data Kemiskinan di Indonesia


Nama Mahasiswa : Tri Hastuti


Nomor Pokok Mahasiswa : 1617031043

Program Studi : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam




Ir. Netti Herawati, M.Sc., Ph.D.
NIP 19650125 199003 2 001


Dr. Aang Nuryaman, S.Si., M.Si.
NIP 19740316 200501 1 001

2. Ketua Jurusan Matematika



Dr. Aang Nuryaman, S.Si., M.Si.
NIP 19740316 200501 1 001

MENGESAHKAN

1. Tim Penguji

Ketua : Ir. Netti Herawati, M.Sc., Ph.D.



Sekretaris : Dr. Aang Nuryaman, S.Si., M.Si.



**Penguji
Utama : Dr. Khoirin Nisa, S.Si., M.Si.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



**Dr. Eng. Supto Dwi Yuwono, S.Si., M.T.
NIP 19740705 200003 1 001**



Tanggal Lulus Ujian Skripsi : 22 Desember 2022

PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertandatangan dibawah ini:

Nama : **Tri Hastuti**
Nomor Pokok Mahasiswa : **1617031043**
Jurusan : **Matematika**
Judul : **Penerapan Model *Geographically Weighted Logistic Regression* (GWLR) dengan Fungsi Pembobot *Adaptive Gaussian Kernel* Pada Data Kemiskinan di Indonesia**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri dan sepanjang pengetahuan saya bukan merupakan hasil yang telah dipublikasikan atau ditulis orang lain atau telah dipergunakan dan diterima sebagai persyaratan penyelesaian studi pada universitas atau institute lain.

Bandar Lampung, 22 Desember 2022



Tri Hastuti
NPM. 1617031043

RIWAYAT HIDUP

Penulis bernama lengkap Tri Hastuti, dilahirkan di Bandar Lampung pada tanggal 27 September 1997 merupakan anak terakhir dari tiga bersaudara, pasangan bapak Sutrisno dan ibu Siti Rohdyah.

Penulis telah menempuh pendidikan di TK Bina Harapan 2003-2004. Kemudian melanjutkan pendidikan di SDN 1 Karang Maritim Bandar Lampung tahun 2004-2010, SMPN 23 Bandar Lampung tahun 2010-2013 dan SMA 10 Bandar Lampung pada tahun 2013-2016.

Pada tahun 2016 penulis terdaftar sebagai mahasiswa S1 Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SNMPTN. Pada tahun 2019 penulis melaksanakan Kerja Praktik (KP) di Badan Pusat Statistik Provinsi Lampung dan melaksanakan Kuliah Kerja Nyata (KKN) di Desa Bakhu, Kecamatan Batu Ketulis, Kabupaten Lampung Barat. Pada tahun 2019 penulis mendapatkan kesempatan untuk mengikuti program Student Exchange di Jepang selama 6 bulan, tepatnya di Universitas *Tokyo University of Agriculture and Technology* (TUAT), Tokyo, Jepang. Selama menjadi mahasiswa penulis aktif di organisasi Himpunan Mahasiswa Matematika (HIMATIKA) FMIPA Unila sebagai anggota bidang Kaderisasi dan Kepemimpinan periode 2017 dan sebagai anggota Persatuan Pelajar Indonesia *Tokyo University of Agriculture and Technology* (PPI TUAT) periode 2019-2020.

PERSEMBAHAN

Puji syukur atas kehadiran Allah SWT atas segala limpahan rahmat dan karunia-Nya sehingga telah memberikan petunjuk dan kemudahan untuk menyelesaikan skripsi ini. Karyaku yang sederhana ini kupersembahkan kepada:

Bapak dan Ibu

Sebagai tanda bakti, hormat dan rasa terimakasihku kepada bapak dan ibu tercinta yang tiada henti untuk mendoakanku, yang telah mendidik dan membesarkanku dengan penuh kasih sayang.

Kakaku Mas Eko dan Mas Yudhi

Terimakasih telah memberikan semangat, do'a dan keceriaan dalam hidupku.

Kakak Iparku Mba Ia dan Ponakanku Khairan

Terima kasih telah memberi keceriaan dalam hidupku.

Dosen Pembimbing dan Penguji

Yang telah berjasa dan senantiasa mengarahkan.

Almamater Tercinta, Universitas Lampung

KATA INSPIRASI

“Allah Pencipta langit dan bumi. Apabila Dia hendak menetapkan sesuatu, Dia hanya berkata “Jadilah!” Maka jadilah sesuatu itu”

(QS. Al-Baqarah :117)

“jalan Allah yang milik-Nya-lah apa yang ada di langit dan apa yang ada di bumi. Ingatlah, segala urusan kembali kepada Allah”

(QS. Asy-Syura :53)

“Wahai orang-orang yang beriman! Mohonlah pertolongan (kepada Allah) dengan sabar dan sholat. Sungguh, Allah bersama orang-orang yang sabar”

(QS. Al-Baqarah :153)

“Raihlah ilmu dan untuk meraih ilmu belajarlah tenang dan sabar”

(Umar bin Khattab)

SANWACANA

Alhamdulillah puji syukur kehadirat Allah SWT atas segala limpahan rahmat dan hidayah-Nya sehingga skripsi ini dapat terselesaikan.

Skripsi dengan judul “Penerapan Model *Geographically Weighted Logistic Regression* (GWLR) dengan Fungsi Pembobot *Adaptive Gaussian Kernel* Pada Data Kemiskinan di Indonesia” merupakan salah satu syarat untuk memperoleh gelar Sarjana Matematika di Universitas Lampung.

Dalam menyelesaikan skripsi ini, banyak pihak yang telah membantu memberikan arahan, bimbingan serta semangat dan dukungan kepada penulis. Oleh karena itu, dengan ketulusan hati penulis ingin mengucapkan terima kasih kepada:

1. Ibu Ir. Netti Herawati, M.Sc., Ph.D., selaku dosen pembimbing satu sekaligus pembimbing akademik yang telah membimbing dan memberikan pengarahan kepada penulis selama perkuliahan hingga dalam menyelesaikan skripsi ini.
2. Bapak Dr. Aang Nuryaman, S.Si., M.Si. selaku pembimbing dua yang telah memberikan pengarahan dan saran dalam proses penyusunan skripsi ini.
3. Ibu Dr. Khoirin Nisa, S.Si., M.Si. selaku dosen pembahas dan penguji skripsi yang telah memberikan pengarahan dan masukan serta saran-saran dalam perbaikan dan penyelesaian skripsi ini.
4. Bapak Dr. Aang Nuryaman, S.Si., M.Si. selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Bapak Dr. Eng. Suropto Dwi Yuwono, S.Si., M.T., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

6. Seluruh dosen dan staff Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Kedua orang tuaku dan kakak yang selalu memberikan do'a dan dukungan tak terhingga serta menjadi penyemangat tersendiri bagi penulis.
8. Teman-teman dekat penulis (Helen, Anan, Lidia, Nabilla, Restu, Risma, Intan, Sipa, Putkar, Feby, Palen dan Keluarga Bakhu) serta orang-orang yang hadir dalam kehidupan perkuliahan penulis yang telah membantu, saling menguatkan, memberikan semangat, do'a dan berbagi keceriaan kepada penulis.
9. Yunna, Hilda, dan Irfan yang telah menjadi tempat bertukar pikiran, memberikan saran dan saling menyemangati selama proses penyelesaian skripsi.
10. Keluarga HIMATIKA terutama anggota bidang kaderisasi dan kepemimpinan periode 2017 serta Teman-teman Jurusan Matematika angkatan 2016 yang telah memberikan pengalaman berharga.
11. Seluruh pihak yang tidak dapat disebutkan satu persatu yang telah membantu serta mendoakan penulis dalam menyelesaikan skripsi ini.

Bandar Lampung, 22 Desember 2022
Penulis,

Tri Hastuti

DAFTAR ISI

	Halaman
DAFTAR TABEL	xiv
I. PENDAHULUAN	1
1.1 Latar Belakang dan Masalah	1
1.2 Tujuan Penelitian.....	3
1.3 Manfaat Penelitian	3
II. TINJAUAN PUSTAKA	4
2.1 Regresi Logistik	4
2.2 Penaksir Parameter Model Regresi Logistik	5
2.3 Pengujian Parameter Model Regresi Logistik.....	6
2.4 Asumsi Model Regresi Logistik.....	7
2.5 Diagnostik Multikolinearitas	8
2.6 Pemilihan <i>Bandwidth</i> dan Model Terbaik.....	8
2.7 Model <i>Geographically Weighted Logistic Regression</i> (GWLR)	9
2.8 Penaksir Parameter GWLR	11
2.9 Pengujian Parameter Model GWLR	12
III. METODOLOGI PENELITIAN	13
3.1 Waktu dan Tempat Penelitian	13
3.2 Data Penelitian	13
3.3 Metode Penelitian.....	15
IV. HASIL DAN PEMBAHASAN	17
4.1 Diagnostik Multikolinearitas pada Data Real	17

4.2 Analisis GWLR Menggunakan Fungsi Pembobot <i>Adaptive Gaussian Kernel</i> pada Data Real.....	18
---	----

V. KESIMPULAN.....	29
---------------------------	-----------

DAFTAR PUSTAKA	30
-----------------------------	-----------

LAMPIRAN

DAFTAR TABEL

Tabel	Halaman
1. Pengelompokan Persentase Tingkat Kemiskinan	13
2. Data Persentase Penduduk Miskin di Indonesia Tahun 2021	14
3. Tingkat Korelasi Antarvariabel Bebas pada Data Real	17
4. <i>Variance Inflation Factor</i> (VIF) pada Data Real.....	17
5. Jarak <i>Euclidean</i> dan Matriks Pembobot Adaptive Gaussian Kernel pada data Real	18
6. Fungsi Logit GWLR dengan Fungsi Pembobot Adaptive Gaussian Kernel pada Data Real	24
7. Pengujian Parameter Model GWLR Adaptive Gaussian Kernel Provinsi Jawa Tengah pada Data Real	26
8. Hasil Pengujian Parameter Model GWLR dengan Fungsi Pembobot Adaptive Gaussian Kernel pada Data Real.....	27

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi adalah suatu metode statistik yang digunakan untuk menentukan hubungan atau pengaruh antara satu variabel terhadap variabel yang lain, yaitu antara variabel respon terhadap variabel prediktor. Variabel respon yang bersifat kategori dianalisis menggunakan analisis regresi logistik. Fotheringham, dkk. (2002) mengatakan segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang lebih dekat akan lebih berpengaruh daripada sesuatu yang jauh. Analisis dengan regresi logistik kurang tepat apabila diterapkan pada data yang dipengaruhi lokasi secara geografis atau biasa disebut dengan data spasial, karena analisis regresi logistik mengabaikan pengaruh dari lokasi tersebut. Pengaruh spasial tidak boleh diabaikan karena akan mengurangi kebaikan model. Oleh karena itu dikembangkan sebuah metode analisis yang dipengaruhi oleh faktor geografis.

Fotheringham, dkk. (2002) mengembangkan sebuah metode untuk menganalisis data spasial dengan memperhatikan faktor geografis yang kemudian diberi nama *Geographically Weighted Regression (GWR)*. Pada analisis regresi variabel respon bersifat kategori, Atkinson, dkk. (2003) menuliskan bahwa GWR dikembangkan untuk memprediksi atau menduga model dari kumpulan data yang memiliki variabel respon biner melalui model logistik. Dan untuk model logistik telah dikembangkan menjadi GWLR. GWLR adalah metode yang merupakan bentuk lokal dari regresi logistik dimana lokasi diperhatikan dan diasumsikan

bahwa data variabel dependen berdistribusi binomial yang digunakan untuk menganalisis data spasial dari proses yang *nonstationer* (Desriwendi, dkk., 2015).

Beberapa penelitian sebelumnya tentang GWLR telah dilakukan diantaranya oleh Atkinson, dkk. (2003) yang mengkaji tentang hubungan antara erosi sungai dengan beberapa variabel yang memengaruhi erosi di sungai Dyfi Afon West Wales. Berdasarkan penelitian Yunida, dkk. (2019) yaitu pendugaan parameter *ordinary least square* (OLS) digunakan pada pemodelan regresi spasial menggunakan GWLR. Penelitian Hasriana, dkk. (2017) yaitu pendugaan parameter *maximum likelihood estimation* (MLE) digunakan pada Pemodelan kemiskinan menggunakan *Geographically Weighted Logistic Regression* dengan Fungsi Pembobot Fixed Kernel”.

Matriks pembobot dapat dibentuk menggunakan suatu fungsi pembobot dimana fungsi tersebut tergantung pada ukuran berketetanggaan (*neighbourhood size*) atau biasa disebut *bandwidth* (Desriwendi, dkk., 2015). Terdapat dua pembobot fungsi kernel yaitu *Fixed Gaussian Kernel* dan *Adaptive Gaussian Kernel*. Pada penelitian Hasriana, dkk. (2017) fungsi pembobot fixed kernel digunakan untuk pemodelan kemiskinan menggunakan GWLR. Fungsi Pembobot *Adaptive Gaussian Kernel* digunakan pada model sesuai dengan kondisi titik pengamatan dan menghasilkan nilai *bandwidth* yang konstan atau berbeda-beda pada setiap lokasi pengamatan. Dalam pengoperasiannya dibutuhkan pemilihan *bandwidth* optimal, karena pemilihan *bandwidth* merupakan langkah penting yang harus dilakukan pada pembobotan. Salah satu metode untuk mencari bandwidth yang optimal adalah dengan menggunakan kriteria *cross validation* (CV).

Kemiskinan merupakan salah satu fenomena keheterogenan spasial, hal ini dapat ditunjukkan dengan kondisi ekonomi yang bervariasi pada masing-masing daerah. Tahun 2021, Indonesia mengalami peningkatan persentase kemiskinan yang cukup tinggi. Mengingat pentingnya pengentasan kemiskinan serta

persentase kemiskinan yang masih tinggi di Indonesia, yaitu sebesar 9,78% pada tahun 2020 dan mengalami peningkatan sebesar 10,14% pada tahun 2021.

Sejauh yang penulis ketahui, analisis terkait kemiskinan di Indonesia yang menggunakan GWLR belum pernah dilakukan, khususnya analisis yang menggunakan *software* R 4.2.1. Maka pada penelitian ini akan menganalisis faktor-faktor yang sangat mempengaruhi persentase kemiskinan di Indonesia, sehingga diperoleh estimasi parameter GWLR dengan Fungsi Pembobot *Adaptive Gaussian Kernel* yang mampu mengidentifikasi faktor-faktor yang sangat mempengaruhi persentase kemiskinan semua provinsi di Indonesia pada tahun 2021

1.2 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah untuk menerapkan model GWLR dengan fungsi pembobot *Adaptive Gaussian Kernel* guna mengetahui faktor-faktor yang mempengaruhi persentase kemiskinan di Indonesia.

1.3 Manfaat Penelitian

Manfaat dari penelitian ini adalah menambah pengetahuan bagi penulis dan memberi masukan kepada para peneliti dan pembaca tentang metode *Geographically Weighted Logistic Regression* (GWLR) dengan pembobot *Adaptive Gaussian Kernel*.

II. TINJAUAN PUSTAKA

2.1 Regresi Logistik

Metode regresi merupakan analisis data yang mendeskripsikan antara sebuah variabel dependen dan satu atau lebih variabel independen. Regresi logistik merupakan metode yang dapat digunakan untuk mencari hubungan antara variabel dependen yang dilambangkan dengan y yang bersifat *dichotomus*, yaitu mempunyai skala nominal dengan dua kategori atau *polychotomous*, yaitu mempunyai skala nominal dengan lebih dari dua kategori, dengan satu atau lebih variabel independen yang dilambangkan dengan x , sedangkan variabel dependennya bersifat kategorik (Agresti, 2002).

Regresi logistik termasuk dalam model linier umum (*Generalized Linear Models* atau GLM). Model linier umum merupakan pengembangan dari model linear klasik. Pada model linier umum komponen acak tidak harus mengikuti distribusi normal, tetapi harus dalam keluarga eksponensial (Yunus, 2016).

Hasil observasi variabel acak dependen (y) mempunyai dua fungsi kategorik yaitu 0 dan 1, sehingga mengikuti distribusi bernoulli dengan fungsi distribusi peluang (Hosmer & Lemeshow, 1989) :

$$P(Y = y) = \pi^y (1 - \pi)^{1-y} ; y = 0 \text{ dan } 1 \quad (2.1)$$

Dimana jika $y = 0$, maka $P(Y = 0) = 1 - \pi$ dan jika $y = 1$, maka $P(Y = 1) = \pi$.

Rata-rata bersyarat dari y jika diberikan nilai x adalah $\pi(x) = E(y|x)$. Pada regresi logistik dapat multivariabel dengan p variabel independen adalah:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \quad (2.2)$$

Dengan mengambil fungsi logit $g(x)$ yaitu :

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p = \beta_0 + \sum_{i=1}^p \beta_i x_i \quad (2.3)$$

Maka model regresi logistik dapat dituliskan dalam bentuk sebagai berikut:

$$\pi(x) = \frac{\exp(g(x))}{1 + \exp(g(x))} \quad (2.4)$$

2.2 Penaksir Parameter Model Regresi Logistik

Untuk menentukan estimasi parameter (β), digunakan metode Maximum Likelihood Estimation (MLE). Parameter (β) diestimasi dengan cara memaksimalkan fungsi *likelihood*. Fungsi *likelihood* pada regresi logistik adalah sebagai berikut (Hosmer & Lemeshow, 2000).

$$\begin{aligned} L(\beta) &= \prod_{i=1}^p \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \\ &= \prod_{i=1}^p \pi(x_i)^{y_i} (1 - \pi(x_i)) \left(\frac{1}{1 - \pi(x_i)} \right)^{y_i} \\ &= \prod_{i=1}^p \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right)^{y_i} (1 - \pi(x_i)) \\ &= \left\{ \prod_{i=1}^p \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right)^{y_i} (1 - \pi(x_i)) \right\} \\ &= \left\{ \prod_{i=1}^p \exp \left[\ln \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right)^{y_i} \right] \right\} \left\{ \prod_{i=1}^p (1 - \pi(x_i)) \right\} \\ &= \left\{ \exp \sum_{i=1}^p y_i \left(\beta_0 + \sum_{j=1}^p \beta_j x_{ij} \right) \right\} \left\{ \prod_{i=1}^p (1 - \pi(x_i)) \right\} \end{aligned}$$

Sesuai dengan persamaan (2.3), maka fungsi *likelihood* adalah

$$l(\beta) = \left\{ \exp \sum_{i=1}^p y_i (\beta_0 + \sum_{j=1}^p \beta_j x_{ij}) \right\} \left\{ \prod_{i=1}^p 1 + \exp(\beta_0 + \sum_{j=1}^p \beta_j x_{ij})^{-1} \right\}$$

Dan persamaan *ln likelihood* yang terbentuk adalah:

$$l(\beta) = \ln l(\beta) = \sum_{i=1}^p y_i (\beta_0 + \sum_{j=1}^p \beta_j x_{ij}) - \sum_{i=1}^p \ln(1 + \exp(\beta_0 + \sum_{j=1}^p \beta_j x_{ij}))$$

Persamaan *ln likelihood* diturunkan terhadap β , untuk mendapatkan nilai β yang dapat memaksimumkan $l(\beta)$. Kemudian hasil yang diperoleh dibuat sama dengan 0.

$$\frac{\partial L(\beta)}{\partial \beta} = 0$$

$$\frac{\partial L(\beta)}{\partial \beta} = \sum_{i=1}^p y_i x_{ij} - \sum_{i=1}^p x_{ij} \left(\frac{\exp(\beta_0 + \sum_{j=1}^p \beta_j x_{ij})}{1 + \exp(\beta_0 + \sum_{j=1}^p \beta_j x_{ij})} \right) = 0$$

Sehingga persamaan *likelihood* adalah:

$$\begin{aligned} \sum_{i=1}^p y_i x_{ij} - \sum_{i=1}^p x_{ij} (\pi(x_i)) &= 0 \\ \sum_{i=1}^p (y_i - \pi(x_i)) x_{ij} &= 0 \end{aligned}$$

Karena persamaan diatas nonlinear maka untuk mendapatkan nilai maksimum *likelihood* maka digunakan metode iteratif yaitu iterasi *newton-raphson*.

2.3 Pengujian Parameter Model Regresi Logistik

Pengujian parameter ini bertujuan untuk menguji apakah variabel prediktor (β) yang terdapat dalam model berpengaruh atau tidak terhadap variabel responnya. Adapun uji yang dilakukan adalah sebagai berikut:

- Untuk menguji signifikansi parameter β dalam model secara bersama-sama dengan menggunakan statistik uji G (Hosmer & Lemeshow, 1989).

Hipotesis:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0.$$

$$H_1: \text{paling sedikit satu } \beta_j \neq 0; j = 1, 2, \dots, p.$$

Taraf signifikansi: α .

Statistik uji:

$$G = -2 \ln \left[\frac{\binom{n-1}{n}^{n_1} \binom{n_0}{n}^{n_0}}{\prod_{i=1}^n (\hat{\pi}(x_i))^{y_i} (1 - (\hat{\pi}(x_i))^{(1-y_i)})} \right] \quad (2.5)$$

dengan $n_1 = \sum_{i=1}^n y_i$; $n_0 = \sum_{i=1}^n (1 - y_i)$; $n = n_0 + n_1$

kriteria uji:

tolak H_0 jika $G > \chi_{(\alpha,p)}^2$ dengan p adalah banyaknya variabel independen.

b. Uji Parsial

Pengujian ini dilakukan untuk mengetahui signifikansi parameter β terhadap variabel respon secara parsial dengan menggunakan statistik uji Wald (Anggraini & purhadi, 2012).

Hipotesis:

$H_0: \beta_j = 0 ; j = 1, 2, \dots, p.$

$H_1: \beta_j \neq 0 ; j = 1, 2, \dots, p.$

Taraf signifikansi: $\alpha.$

Statistik uji:

$$Z_j = \frac{\hat{\beta}_j}{SE\hat{\beta}_j} \text{ atau } W_j^2 = \frac{\hat{\beta}_j^2}{(SE(\hat{\beta}_j))^2} \quad (2.6)$$

Dimana, $SE(\hat{\beta}_j) = \sqrt{var(\hat{\beta}_j)}$.

Kriteria uji:

tolak H_0 jika $|Z_j| > Z_{\alpha/2}$ atau tolak H_0 jika $W_j^2 > \chi_{\alpha,1}^2$.

2.4 Asumsi Model Regresi Logistik

Regresi logistik tidak memerlukan asumsi normalitas, heteroskedasitas, dan autokorelasi, dikarenakan variabel respon yang terdapat pada regresi logistik merupakan variabel *dummy* (0 dan 1), sehingga residualnya tidak memerlukan pengujian tersebut (Caraka & Yasin, 2017). Untuk asumsi multikolinearitas, karena hanya melibatkan variabel – variabel prediktor , maka masih perlu untuk dilakukan pengujian. Untuk pengujian multikolinearitas ini dapat dilakukan uji kebaikan sesuai (*goodness of fit test*), yang kemudian dilanjutkan dengan pengujian parameter, guna melihat variabel-variabel prediktor mana saja yang signifikan, sehingga dapat tetap digunakan dalam penelitian.

2.5 Diagnostik Multikolinearitas

Salah satu asumsi yang harus dipenuhi pada analisis regresi dengan beberapa peubah prediktor adalah tidak adanya korelasi antara satu peubah prediktor dengan peubah prediktor yang lain. Menurut Sarwoko (2005), ada beberapa metode untuk mendeteksi multikolinieritas diantaranya:

- a. Kolinieritas seringkali diduga ketika R^2 tinggi (misalnya: antara 0,7 dan 1) dan jika R^2 tinggi, ini berarti bahwa uji F dari prosedur analisis varian dalam sebagian kasus akan menolak hipotesis nol bahwa nilai koefisien kemiringan parsial secara simultan sebenarnya adalah nol.
- b. Multikolinieritas timbul karena satu atau lebih variabel yang menjelaskan merupakan kombinasi liner yang mendekati pasti dari variabel yang menjelaskan lainnya.

Menurut Montgomery & Runger (2011), multikolinearitas dapat dideteksi menggunakan nilai *Variance Inflation Factor* (VIF). Nilai VIF dapat dicari menggunakan rumus sebagai berikut:

$$VIF_j = \frac{1}{(1-R_j^2)} \quad j = 1, 2, \dots, k \quad (2.7)$$

Dimana, R_j^2 merupakan koefisien determinasi yang didapat dari variabel predictor X_j yang diregresikan dengan variabel predictor lainnya. Jika X_j tidak berkorelasi dengan peubah bebas lain, maka R_j^2 akan bernilai kecil dan nilai VIF akan mendekati 1. Sebaiknya jika X_j mempunyai korelasi dengan peubah bebas lain, maka R_j^2 akan mendekati 1 dan nilai VIF menjadi besar. Jika nilai VIF lebih besar dari 10, maka menunjukkan adanya multikolinearitas (Montgomery & Peck, 1992).

2.6 Pemilihan *Bandwidth* dan Model Terbaik

Menurut Fotheringham, dkk. (2002) pemilihan ukuran *bandwidth* menjadi salah satu hal yang penting karena akan mempengaruhi ketepatan hasil regresi.

Untuk mencari pembobot pada masing-masing lokasi didasarkan pada jarak *euclidean* d_{ij} dan *bandwidth* (h) yang dihasilkan pada masing-masing lokasi. Pada penelitian ini digunakan metode *Cross Validation* (CV) dengan rumus sebagai berikut:

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(h)]^2 \quad (2.9)$$

Dengan $\hat{y}_{\neq i}(h)$ adalah nilai penaksir y_i dimana pengamatan di lokasi (u_i, v_i) dihilangkan dari proses estimasi. Untuk mendapatkan nilai *bandwidth* (h) yang optimal maka diperoleh dari h yang menghasilkan CV yang minimum (Caraka & Yasin, 2017). Untuk pemilihan model terbaik ditentukan dengan nilai *Akaike Information Criterion* (AIC) yang paling kecil dengan persamaan berikut:

$$AIC = 2n \log_e(\hat{\sigma}) + n \log_e(2\pi) + n + tr(\mathbf{S}) \quad (2.10)$$

2.7 Model Geographically Weighted Logistic Regression (GWLR)

GWLR adalah salah-satu metode untuk mendapatkan parameter regresi dengan memperhitungkan faktor spasial dan merupakan pendekatan alternatif dari regresi logistik yang menggabungkan parameter nonstasioner dan data kategorikal. Dalam model GWLR, variabel dependen diprediksi dengan variabel independen yang masing-masing koefisien regresinya bergantung pada lokasi dimana data tersebut diamati. Model GWLR dapat ditulis sebagai berikut:

$$\pi(x_i) = \frac{\exp(\sum_{k=0}^p \beta_k(u_i, v_i) x_{ik})}{1 + \exp(\sum_{k=0}^p \beta_k(u_i, v_i) x_{ik})}; \quad i = 1, 2, \dots, n \quad (2.11)$$

dengan,

- $\pi(x_i)$ = nilai observasi variabel respon ke- i ;
- $\beta_k(u_i, v_i)$ = vektor nilai observasi variabel prediktor ke- k pada lokasi pengamatan ke- i ;
- (u_i, v_i) = menyatakan koordinat letak geografis (*longitude*, *latitude*) dari lokasi pengamatan ke- i ;
- x_{ik} = nilai observasi variabel prediktor ke- k pada lokasi pengamatan ke- i .

Dengan demikian setiap parameter dihitung pada setiap titik lokasi geografis. Hal ini menghasilkan nilai variasi pada setiap parameter regresi di suatu kumpulan wilayah geografis. Jika nilai parameter regresi konstan pada tiap-tiap wilayah geografis maka model GWLR adalah model global. Artinya, tiap-tiap geografis mempunyai model yang sama. Hal ini merupakan kasus khusus dari GWLR. Peran pembobot sangatlah penting, karena nilai pembobot ini mewakili letak data observasi satu dengan lainnya. Oleh karena itu, sangat dibutuhkan ketepatan pembobot. Pada analisis spasial, diperlukan pembobot spasial pada masing-masing lokasi ke- i . Apabila lokasi ke- j terletak pada koordinat (u_j, v_j) maka akan diperoleh jarak *Euclidean* antara lokasi ke- i dan lokasi ke- j dengan menggunakan persamaan:

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (2.12)$$

dengan,

- d_{ij} = Jarak *Euclidean* lokasi pengamatan ke- i dengan lokasi pengamatan ke- j
- u_i = Longitude jarak lokasi pengamatan i
- u_j = Longitude jarak lokasi pengamatan j
- v_i = Latitude jarak lokasi pengamatan i
- v_j = Latitude jarak lokasi pengamatan j

Menurut Fotheringham, dkk. (2002) jika terdapat nilai pengamatan ke- j yang jaraknya terlalu jauh dari lokasi i maka pengamatan yang jaraknya diluar radius r dari lokasi i dihilangkan, yaitu dengan memberikan nilai nol untuk pembobot pada pengamatan yang jaraknya lebih besar dari r . Penggunaan metode *adaptive Kernel* bagi penentuan nilai *bandwidth* akan cocok untuk pengamatan-pengamatan yang tersebar dengan pola yang tidak beraturan dan terkelompok. Metode *adaptive Kernel* memungkinkan untuk mendapatkan nilai *bandwidth* yang berbeda-beda untuk semua titik pengamatan. Hal ini disebabkan karena metode *adaptive Kernel* dapat menyesuaikan dengan kondisi titik pengamatan.

Pembobot ini dapat ditulis:

$$w_j(u_i, v_i) = \begin{cases} 1, & \text{jika } d_{ij} \leq r \\ 0, & \text{jika } d_{ij} > r \end{cases} \quad (2.13)$$

Metode yang digunakan adalah *Adaptive Gaussian Kernel* sebagai berikut:

$$w_{ij} = \exp \left[- \left(\left(d_{ij} / h_{i(q)} \right) / 2 \right)^2 \right] \quad (2.14)$$

Dengan w_{ij} adalah matriks pembobot lokasi pengamatan ke- i dengan lokasi pengamatan ke- j , h adalah parameter penghalus (*bandwidth*) dan $h_{i(q)}$ adalah *bandwidth* adaptif atau *bandwidth* yang berbeda untuk setiap lokasi yang menetapkan q sebagai jarak tetangga terdekat (*nearest neighbor*) dari lokasi i (Fotheringham, dkk. 2002).

2.8 Penaksir Parameter GWLR

Model GWLR merupakan pengembangan dari model regresi logistik yang menghasilkan bentuk penaksir parameter model bersifat lokal untuk setiap lokasi tempat data tersebut didapatkan. Penaksiran parameter model GWLR dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Sesuai dengan persamaan diatas maka fungsi *likelihood* yang terbentuk adalah:

$$L(\beta(u_i, v_i)) = \left\{ \exp \sum_{i=1}^n y_i (\beta_0(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i) x_{ij}) \right\} \left\{ \prod_{i=1}^n (1 + \exp(\beta_0(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i) x_{ij}))^{-1} \right\}$$

$$\ln L(\beta(u_i, v_i)) = \sum_{i=1}^n y_i (\beta_0(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i) x_{ij}) - \sum_{i=1}^n \ln(1 + \exp(\beta_0(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i) x_{ij}))$$

Letak geografis merupakan pembobot pada model ini, sehingga pembobot diberikan dalam bentuk:

$$\ln L(\beta(u_i, v_i)) = \sum_{i=1}^n y_i w_{ij}(u_i, v_i) (\beta_0(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i) x_{ij}) - \sum_{i=1}^n w_{ij}(u_i, v_i) \ln(1 + \exp(\beta_0(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i) x_{ij}))$$

Persamaan diatas diturunkan terhadap $\beta_j(u_i, v_i)$, agar mendapatkan nilai β yang dapat memaksimumkan. Hasil yang diperoleh disamadengankan 0. Selanjutnya, karena hasil dari turunan partial bersifat *nonlinear*, maka digunakan metode iteratif *Newton-Rhapson*.

Dan bentuk umum dari hasil penurunannya adalah:

$$\beta^{t+1}(u_i, v_i) = \beta^{(t)}(u_i, v_i) - H^{(t)-1}(\beta^{(t)}(u_i, v_i)) g^{(t)}(\beta^{(t)}(u_i, v_i))$$

Iterasi berhenti pada saat konvergen, yaitu saat:

$$\|\beta^{(t+1)}(u_i, v_i) - \beta^{(t)}(u_i, v_i)\| \leq \varepsilon$$

2.9 Pengujian Parameter Model GWLR

Menguji parameter secara parsial yang bertujuan untuk mengetahui parameter mana saja yang signifikan mempengaruhi variabel dependennya.

Hipotesis:

$$H_0 : \beta_k(u_i, v_i) = 0$$

$$H_1 : \beta_k(u_i, v_i) \neq 0 ; k = 1, 2, \dots, p ; i = 1, 2, \dots, n$$

Statistik uji Wald:

$$W = \frac{\hat{\beta}_k(u_i, v_i)}{SE(\hat{\beta}_k(u_i, v_i))} \quad (2.15)$$

dengan $SE(\hat{\beta}_k(u_i, v_i)) = \sqrt{var(\hat{\beta}_k(u_i, v_i))}$.

Kriteria uji yaitu tolak H_0 jika $|W| > Z_{\alpha/2}$

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun akademik 2022-2023 dan bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang didapat dari Badan Pusat Statistik (BPS). Data yang digunakan adalah persentase penduduk miskin di Indonesia (Y). Sedangkan variabel independennya (X) yaitu pengeluaran per kapita (X_1), rata-rata lama sekolah (X_2), Produk Domestik Regional Bruto (PDRB) per kapita (X_3), kepadatan penduduk (X_4). Menurut BPS, rata-rata persentase penduduk miskin di Indonesia pada tahun 2021 sebesar 10,14%. Persentase penduduk miskin di Indonesia dikategorikan menjadi dua yaitu kategori miskin dan tidak miskin.

Tabel 1. Pengelompokan Persentase Tingkat Kemiskinan

Keterangan	Persentase Tingkat Kesmiskinan	Kategori
Tidak Miskin	< 10,14%	0
Miskin	> 10,14%	1

Dengan faktor-faktor yang mempengaruhi persentase tingkat kemiskinan adalah pengeluaran per kapita, rata-rata lama sekolah, dan produk domestik regional bruto (PDRB) per kapita, kepadatan penduduk. Data persentase penduduk miskin di Indonesia tahun 2021 dapat dilihat pada tabel 2 berikut.

Tabel 2. Data Persentase Penduduk Miskin di Indonesia Tahun 2021

Wilayah	Y	X1	X2	X3	X4
ACEH	1	9572	9.37	34680.46	92
SUMATERA UTARA	0	10499	9.58	57569.79	205
SUMATERA BARAT	0	10790	9.07	45293.75	133
RIAU	0	10736	9.19	129852.6	75
JAMBI	0	10588	8.6	65193.22	72
SUMATERA SELATAN	1	10662	8.3	57487.44	93
BENGKULU	1	10487	8.87	39143.43	102
LAMPUNG	1	10038	8.08	40950.42	262
KEP. BANGKA BELITUNG	0	12819	8.08	58338.82	90
KEP. RIAU	0	14122	10.18	130125.2	258
DKI JAKARTA	0	18520	11.17	274709.6	15978
JAWA BARAT	0	10934	8.61	45299.58	1379
JAWA TENGAH	1	11034	7.75	38669.11	1120
DI YOGYAKARTA	1	14111	9.64	40229.83	1185
JAWA TIMUR	1	11707	7.88	60043.33	855
BANTEN	0	12033	8.93	55210.65	1248
BALI	0	13820	9.06	50381.21	755
NUSA TENGGARA BARAT	1	10377	7.38	26002.48	290
NUSA TENGGARA TIMUR	1	7554	7.69	2058113	111
KALIMANTAN BARAT	0	8974	7.45	42282.9	37
KALIMANTAN TENGAH	0	11182	8.64	62912.85	18
KALIMANTAN SELATAN	0	12143	8.34	46712.68	106
KALIMANTAN TIMUR	0	12116	9.84	182540.8	30
KALIMANTAN UTARA	0	9075	9.11	155080.6	9
SULAWESI UTARA	0	10882	9.62	54043.18	190
SULAWESI TENGAH	1	9378	8.89	81733.04	49
SULAWESI SELATAN	0	11184	8.46	59656.24	196
SULAWESI TENGGARA	1	9381	9.13	52293.97	70
GORONTALO	1	10157	7.9	37170.45	105
SULAWESI BARAT	1	9153	7.96	35036.02	86

Tabel 2. Lanjutan

Wilayah	Y	X1	X2	X3	X4
MALUKU	1	8770	10.03	26072.98	40
MALUKU UTARA	0	8140	9.09	40302.32	41
PAPUA BARAT	1	7929	7.69	73539	11
PAPUA	1	6955	6.76	54034.26	14

dengan,

- Y = Persentase Penduduk Miskin di Indonesia;
- X_1 = Pengeluaran Per Kapita (ribu rupiah);
- X_2 = Rata-rata lama sekolah (tahun);
- X_3 = Produk Domestik Regional Bruto (PDRB) Per Kapita
- X_4 = Kepadatan Penduduk (jiwa).

3.3 Metode Penelitian

Penelitian ini dilakukan dengan mempelajari buku-buku teks penunjang dan karya ilmiah yang disajikan dalam bentuk jurnal. Penulis menggunakan *software* R versi 4.2.1 untuk mempermudah perhitungan dengan hasil yang akurat.

Adapun langkah-langkah yang dilakukan adalah sebagai berikut:

1. Memeriksa kolinieritas dari variabel-variabel independen dengan melihat matriks kolerasi dan nilai *Variance Inflation Factor* (VIF) dari data yang didapat.
2. Menentukan nilai *bandwidth* menggunakan metode *Cross Validation* (CV) dari data yang didapat, dengan rumus sebagai berikut:

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(h)]^2 \quad (3.1)$$

dengan $\hat{y}_{\neq i}(h)$ adalah nilai penaksir y_i dimana pengamatan di lokasi (u_i, v_i) dihilangkan dari proses estimasi.

3. Menghitung jarak *Euclidean* antarlokasi pengamatan ke- i dengan lokasi pengamatan ke- j berdasarkan letak geografis untuk setiap provinsi, dengan rumus sebagai berikut:

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (3.2)$$

4. Menghitung matriks pembobot menggunakan fungsi *Adaptive Gaussian Kernel* dengan *bandwidth* yang sama pada setiap lokasi untuk data yang didapat dengan rumus sebagai berikut:

$$w_{ij} = \exp \left[- \left(\left(d_{ij} / h_{i(q)} \right) / 2 \right)^2 \right] \quad (3.3)$$

5. Menganalisis penaksir parameter model GWLR dengan fungsi pembobot *Adaptive Gaussian Kernel*.
6. Membuat kesimpulan.

V. KESIMPULAN

Berdasarkan hasil dan pembahasan yang sudah dilakukan, maka diperoleh kesimpulan bahwa model GWLR dengan fungsi pembobot *Adaptive Gaussian Kernel* menghasilkan parameter yang bersifat lokal disetiap titik atau wilayah dimana data tersebut diamati. Maka itu didapatkan model GWLR dengan fungsi pembobot *Adaptive Gaussian Kernel* tiap Provinsi di Indonesia dan variabel independen yang berpengaruh secara signifikan terhadap persentase kemiskinan di Indonesia tahun 2021 yaitu Produk Domestik Regional Bruto (PDRB) per kapita (X_3).

DAFTAR PUSTAKA

- Agresti, A. 2002. *Categorical Data Analysis*. 2nd Edition. John Wiley & Sons, New York.
- Anggarini & Purhadi. 2012. Pemodelan Faktor-faktor yang Berpengaruh Terhadap Prevalensi Balita Kurang Gizi di Provinsi Jawa Timur dengan Pendekatan Geographically Weighted Logistic Regression (GWLR). *Jurnal Sains dan Seni ITS*. 1(1): 47-52.
- Atkinson, P.M., German, S.E., Sear, D.A., & Clark, M.J. 2003. *Exploring The Relations Between Riverbank Erosion and Geomorphological Controls Using Geographically Weighted Logistic Regression*. Ohio State University, Ohio.
- Caraka, R.E. & Yasin, H. 2017. *Geographically Weighted Regression (GWR) Sebuah Pendekan Regresi Geografis*. Mobius, Yogyakarta.
- Desriwendi, Hoyyi, A., & Wuryandari, T. 2015. Pemodelan Geographically Weighted Logistic Regression (GWLR) dengan Fungsi Pembobot Fixed Gaussian Kernel dan Adaptive Gaussian Kernel. *Jurnal Gaussian*. 4(2):193-204.
- Fotheringham, A.S., Brundson, C., & Charlton, M. 2002. *Geographically Weighted Regression: Analysis of Spatially Varying Relationship*. John Wiley & Sons, England.
- Hasriana, Raupong, & Nirwan, I. 2017. Pemodelan Kemiskinan Menggunakan Geographically Weighted Logistic Regression dengan Fungsi Pembobot Fixed Kernel. Skripsi. Universitas Hasanuddin, Makasar.

Hosmer, D.W. & Lemeshow, S. 1989. *Applied Logistic Regression*. John Wiley & Sons, New York.

Hosmer, D.W. & Lemeshow, S. 2000. *Applied Logistic Regression*. 2nd Edition. Wiley, New York.

Montgomery, D.C. & Peck, A.E. 1992. *Introduction to Linear Regression Analysis*. A Wiley Intersection Publication, New York.

Montgomery, D.C. & Runger, G.C. 2011. *Applied Statistics and Probability For Engineers*. 5th Edition. John Wiley & Sons, New York.

Sarwoko. 2005. *Dasar-dasar Ekonometrika*. Andi, Yogyakarta.

Yunida, D.P., Scolastika, M., & Putriasi, H. 2019. *Pemodelan Regresi Spasial Menggunakan Geographically Weighted Regression*. Skripsi. Univertas Negeri Semarang, Semarang.

Yunus, N.H. 2016. *Estimasi Parameter Model Geographically Weighted Logistic Regression (GWLR) pada Data yang Tidak Mengandung Multikolinieritas*. Skripsi. Universitas Negeri Maulana Malik Ibrahim Malang, Malang.