

Pemodelan Faktor-Faktor Yang Berpengaruh Terhadap Angka Buta Huruf Di Provinsi Sulawesi Selatan Dengan *Geographically Weighted Logistic Regression* (GWLR).

Nurul Era Natasyah*, Muhammad Nusrang, Zakiyah Mar'ah

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia

Keywords: GWLR, Adaptive Gaussian Kernel, Adaptive Bisquare Kernel, Adaptive Tricube Kernel, Illiteracy.

Abstract:

Geographically Weighted Logistic Regression (GWLR) is the development of a logistic regression model applied to spatial data from non-stationary processes with categorical response variables. The high rate of illiteracy is one of the crucial problems in the field of education that has not been resolved to date. South Sulawesi is the 4th province with the highest percentage of illiteracy in Indonesia in 2022. This research aims to obtain the GWLR model and the factors that have a significant influence on the illiteracy rate in South Sulawesi in 2022. In this research, we compare three functions Kernel weightings are Adaptive Gaussian Kernel, Adaptive Bisquare Kernel, and Adaptive Tricube Kernel. Selection of the best model uses the smallest AIC value. The results of this research are that the GWLR model with the Adaptive Tricube Kernel weighting function is the best model in modeling cases of illiteracy in South Sulawesi in 2022 which is obtained based on the smallest AIC value and the factor that has a significant influence on the illiteracy rate is the Open Unemployment Rate (X_1), percentage of poor population (X_2), Elementary School Enrollment Rate (X_3), and area with city status (X_4).

1. Pendahuluan

Analisis regresi adalah suatu metode statistik yang bertujuan untuk mengetahui hubungan dari suatu variabel respon Y terhadap variabel yang lain X , yang disebut variabel prediktor (Tiro, 2010). Model regresi biasanya memiliki variabel respon bersifat kontinu atau diskrit, jika variabel respon yang bersifat diskrit atau kategori dapat dianalisis menggunakan analisis regresi logistik (Purnamasari, 2016).

Model regresi logistik pada umumnya digunakan ketika unit pengamatannya bukan merupakan wilayah atau lokasi, karena model tersebut tidak memperhitungkan efek wilayah atau lokasi yang selanjutnya disebut efek spasial (Pravitasary dkk., 2015). Metode statistika yang digunakan ketika terjadi masalah heterogenitas spasial yaitu apabila satu variabel prediktor yang sama memberikan respon yang berbeda pada setiap lokasi adalah *Geographically Weighted Regression* (GWR). Melalui GWR dapat diketahui pula variasi spasial dalam nilai penaksiran parameter, sehingga akan dihasilkan nilai parameter untuk setiap titik atau lokasi dimana data tersebut diamati (Fotheringham, dkk., 2002). Pemodelan GWR pada data dependen dikotomus yang memuat heterogenitas spasial biasa disebut dengan *Geographically Weighted Logistic Regression* (GWLR) (Sumarni, dkk., 2020). 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 non stasioner (Lestari, 2020).

* Corresponding author.

E-mail address: tasyawtp098@gmail.com



Pendidikan merupakan suatu elemen yang sangat penting dalam perkembangan suatu bangsa. Dengan pendidikan, anak-anak diasah melalui pengetahuan yang positif dalam menemukan dan merumuskan tujuan untuk dirinya di masa-masa mendatang sesuai dengan tujuan pendidikan nasional yang ditetapkan undang-undang nomor 20 tahun 2003 tentang Sistem Pendidikan Nasional. Dalam undang-undang tersebut ditegaskan pendidikan nasional bertujuan mencerdaskan kehidupan bangsa dan meningkatkan manusia Indonesia seutuhnya yaitu manusia yang beriman dan bertakwa terhadap Tuhan Yang Maha Esa, berbudi pekerti luhur (Sisdiknas, 2003).

Melalui pendidikan, manusia akan mendapatkan pengetahuan yang berguna untuk mengenali dan mengembangkan potensi yang dimiliki secara optimal (Muhiyatul, 2016). Kemampuan membaca dan menulis merupakan salah satu kemampuan minimal yang harus dikuasai dan dikembangkan oleh manusia untuk meraih kehidupan yang lebih sejahtera. Meskipun demikian, masih terdapat sejumlah masalah pada bidang pendidikan, salah satu indikatornya adalah angka buta huruf (Satvika, dkk., 2023).

Secara nasional, Sulawesi Selatan bahkan masuk dalam kategori daerah dengan angka buta huruf yang masih tinggi. ABH penduduk usia 15-45 tahun di Sulawesi Selatan 1 dekade terakhir menunjukkan penurunan hingga tahun 2021 akan tetapi data terbaru BPS 2022, untuk usia 15-44 tahun, ABH di Provinsi Sulawesi Selatan naik dari 1,71% pada tahun 2021 menjadi 1,77% pada tahun 2022. Berdasarkan data Badan Pusat Statistika (2022) menyatakan bahwa penduduk di Provinsi Sulawesi Selatan dengan usia 15 tahun ke atas mencapai 6,69%. ABH di provinsi Sulawesi Selatan masih tergolong tinggi karena melebihi angka buta huruf nasional. Dengan tingginya angka buta huruf di Sulawesi Selatan, maka ingin dilakukan penelitian mengenai model regresi yang mampu menangkap efek lokal, dimana variabel respon merupakan data biner melalui metode GWLR. Metode GWLR ini akan diterapkan pada faktor-faktor yang sangat berpengaruh angka buta huruf kabupaten/kota di Sulawesi Selatan.

Penelitian-penelitian sebelumnya tentang pemodelan GWLR yaitu penelitian oleh (Pravitasary dkk., 2015) yaitu Pemodelan Faktor-Faktor yang Berpengaruh Terhadap Angka Buta Huruf di Provinsi Jawa Barat dengan *Geographically Weighted Logistic Regression* menunjukkan bahwa model GWLR lebih baik daripada model regresi logistik global berdasarkan nilai AIC yang diperoleh. Penelitian oleh (Sasmita, 2022) yaitu *Geographically Weighted Logistic Regression (GWLR)* pada Pemodelan Faktor-Faktor yang Mempengaruhi Laju Pertumbuhan Penduduk Kabupaten/Kota di Provinsi Sulawesi Selatan. Hasil penelitian ini menunjukkan bahwa model GWLR lebih baik daripada model regresi logistik berdasarkan nilai AIC yang diperoleh. Penelitian oleh (Ulhaq, 2020) yaitu *Geographically Weighted Logistic Regression (GWLR) with Gaussian Adaptive Kernel Weighting Function, Bisquare, and Tricube In Case Of Malnutrition Of Toddlers In Indonesia* menunjukkan bahwa model terbaik dihasilkan menggunakan model GWLR. Penelitian oleh (Astuti, 2017) yaitu Pemodelan Angka Buta Huruf di Kabupaten/kota se-Jawa Timur dengan Metode *Geographically Weighted t Regression*. Hasil dari penelitian ini menunjukkan bahwa variabel angka partisipasi murni tingkat SD, rasio murid-guru, tingkat pengangguran terbuka dan persentase penduduk miskin berpengaruh signifikan terhadap Angka Buta Huruf di Kabupaten/kota se-Jawa Timur.

Berdasarkan uraian di atas maka akan dilakukan penelitian yang dengan judul “Pemodelan Faktor-Faktor yang Berpengaruh terhadap Angka Buta Huruf di Provinsi Sulawesi Selatan dengan *Geographically Weighted Logistic Regression (GWLR)*”.

2. Tinjauan Pustaka

2.1. Model Regresi Logistik

Model regresi logistik yang variabel responnya berskala kategori biner atau memiliki dua kategori bernilai 0 dan 1 disebut dengan regresi logistik biner (Agresti, 2013). Karena hasil observasi variabel respon memiliki dua kategori, maka mengikuti distribusi Bernoulli dengan fungsi kepadatan peluang sebagai berikut (Hosmer dan Lemeshow, 1989):

$$P(Y = y) = \pi^y(1 - \pi)^{1-y} \text{ dimana } y = 0, 1 \quad (1)$$

Rata-rata bersyarat dari y , apabila nilai x adalah $\pi(x) = E(y|x)$. Model regresi logistik berganda berfungsi jika jumlah variabel prediktor yang digunakan pada regresi logistik lebih dari satu. Maka model regresi logistik dengan k variabel yaitu (Hosmer dan Lemeshow, 2000) :

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)} \quad (2)$$

dengan $\pi(x)$ adalah peluang kejadian sukses dengan nilai probabilitas $0 \leq \pi(x) \leq 1$, β_0 merupakan intercept (bilangan konstant), β_1, \dots, β_p adalah parameter regresi logistic dan x_1, \dots, x_k adalah nilai peubah bebas. Maka transformasi logit pada $\pi(x)$ menjadi:

$$\text{logit}(\pi(x)) = \log \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k \quad (3)$$

Penaksir parameter-parameter dalam model logistik dapat dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Pada dasarnya metode maksimum *likelihood* memberikan nilai penaksir β dengan memaksimalkan fungsi *likelihood* (Hosmer dan Lemeshow, 2000). Oleh karena model regresi logistik merupakan model non linear maka, penaksir kemungkinan maksimum harus diselesaikan secara numerik melalui metode Newton Rhapson.

2.2. Multikolinieritas

Menurut Gujarati (2003), terjadinya hubungan linier antar variabel prediktor dalam suatu model regresi disebut multikolinieritas. Cara mendeteksi multikolinieritas salah satunya dengan melihat nilai VIF (*Varian Inflation Factor*). Nilai VIF lebih besar dari 10 menunjukkan adanya multikolinieritas. Persamaan VIF dinyatakan sebagai berikut (DeMaris, 2004).

$$VIF_k = \frac{1}{1 - R_k^2} \quad (4)$$

dengan R_k^2 adalah koefisien determinasi antara variabel prediktor X_k dengan variabel prediktor lainnya.

2.3. Efek Spasial

Data spasial memiliki karakteristik yang berbeda di setiap lokasi pengamatannya tetapi memiliki hubungan yang cukup erat dengan data di lokasi pengamatan yang berdekatan, hubungan ini disebut efek spasial. Efek spasial yang terjadi antar wilayah pengamatan dapat dikategorikan menjadi ketergantungan spasial (*dependence spasial*) dan keragaman (heterogenitas spasial) (Anselin, 2003).

2.3.1. Dependensi Spasial

Pengujian dependensi spasial dapat dilakukan dengan uji *Moran's I* (Anselin, 2003).

Hipotesis uji *Moran's I* adalah sebagai berikut:

$H_0 : \mu_i = 0$ (tidak ada dependensi antar lokasi)

$H_1 : \mu_i \neq 0$ (ada dependensi antar lokasi)

Statistik Uji

$$Z(I) = \frac{I - E(I)}{\sqrt{\text{Var}(I)}} \quad (5)$$

Dengan

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n c(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i \neq j} w_{ij} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (6)$$

$$E(I) = -\frac{1}{n-1} \quad (7)$$

$$\text{Var}(I) = \frac{n^2 S_1 - n S_2 + 3 S_0^2}{(n^2 - 1) S_0^2} - [E(I)]^2 \quad (8)$$

Keterangan :

x_i : nilai pengamatan pada lokasi ke- i

x_j : nilai pengamatan pada lokasi ke- j

\bar{x} : rata-rata

w_{ij} : matriks pembobot terstandarisasi antara lokasi i dan j

n : banyaknya lokasi pengamatan

Kriteria Uji :

Tolak H_0 , jika $|Z_{hitung}| > Z_{\left(\frac{\alpha}{2}\right)}$ atau $p\text{-value} < \alpha$ yang artinya terdapat dependensi spasial.

2.3.2. Heterogenitas Spasial

Dalam menunjukkan apakah ada heterogenitas atau keragaman daerah dalam spasial, maka digunakan uji *Breusch-Pagan* (Anselin, 2003).

Hipotesis yang diuji adalah :

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_p^2 = \sigma^2$

$H_1 : \text{Minimal ada satu } \sigma_1^2 \neq \sigma^2 : k = 1, 2, \dots, p$

Nilai dari uji *Breusch-Pagan* (BP), yaitu:

$$BP = \frac{1}{2} f^T Z (Z^T Z)^{-1} Z^T f \quad (9)$$

Keterangan :

e_i^2 : galat observasi ke- i ; $i = 1, 2, \dots, n$

$f : (f_1, f_2, \dots, f_n)^T$ dengan $f_i = \left(\frac{e_i^2}{\sigma^2} - 1\right)$

Z : matriks berukuran $n \times (p+1)$ yang berisi vector yang distandarkan untuk tiap observasi

2.4. Geographically Weighted Regression (GWR)

Model GWR adalah pengembangan dari model regresi dimana parameter dihitung pada setiap lokasi pengamatan, sehingga setiap lokasi pengamatan mempunyai nilai parameter yang berbeda-beda (Fotheringham, dkk., 2002). Model GWR menurut Fotheringham, dkk., (2002) dapat dinyatakan pada persamaan :

$$y_i = \beta_o(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (10)$$

dengan nilai y_i adalah nilai observasi variabel respon pada pengamatan ke- i , $\beta_o(u_i, v_i)$ adalah nilai konstanta (*intercept*) pada pengamatan ke- i , $\beta_k(u_i, v_i)$ adalah koefisien regresi variabel prediktor ke- k pada pengamatan ke- i , x_{ik} adalah nilai observasi variabel prediktor ke- k pada pengamatan ke- i , dan ε_i adalah residual ke- i .

2.5. Fungsi Pembobot Spasial

Fungsi pembobot yaitu berperan penting dalam melakukan penaksiran parameter. Titik-titik yang dekat dengan suatu lokasi i akan lebih mempengaruhi penaksiran parameter di suatu lokasi tersebut (Leung, dkk., 2000). Dalam menentukan pembobot maka dilakukan penghitungan dengan cara lokasi j terdapat pada koordinat (u_i, v_i) akan didapatkan jarak *Euclidean* antara i dan lokasi j . perumusan jarak *Euclidean* adalah sebagai berikut (Caraka, dkk., 2017):

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (11)$$

Fungsi *Kernel* merupakan metode yang bisa digunakan, jenis-jenisnya yaitu:

a. Fungsi *Adaptive Gaussian Kernel*

$$w_i(u_i, v_i) = \exp \left[- \left(\frac{d_{ij}/h(q)}{2} \right)^2 \right] \quad (12)$$

b. Fungsi Adaptive Bi-square Kernel

$$w_i(u_i, v_i) = \begin{cases} \left[1 - \left(\frac{d_{ij}}{h_{(q)}}\right)^2\right]^2, & \text{jika } d_{ij} \leq h_{(q)} \\ 0, & \text{jika } d_{ij} > h_{(q)} \end{cases} \quad (13)$$

c. Fungsi Adaptive Tricube Kernel

$$w_i(u_i, v_i) = \begin{cases} \left[1 - \left(\frac{d_{ij}}{h_{(q)}}\right)^3\right]^3, & \text{jika } d_{ij} \leq h_{(q)} \\ 0, & \text{jika } d_{ij} > h_{(q)} \end{cases} \quad (14)$$

Dalam pencarian pembobot lokasi berdasarkan jarak *euclidean* d_{ij} dan *bandwidth*, dimana *bandwidth* (h) merupakan parameter non negative dan disebut dengan parameter penghalus. *Bandwidth* optimum dipilih melalui pendekatan *least square*, dimana pendekatan ini menggunakan metode *Cross Validation* (CV). Rumus CV yaitu :

$$CV(h) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(h))^2 \quad (15)$$

Dimana $\hat{y}_{\neq i}(h)$ merupakan nilai penaksir dari y_i dimana pengamatan di lokasi u_i, v_i selama proses estimasi dihilangkan. Nilai *bandwidth* optimum dihasilkan dari nilai h yang menghasilkan nilai CV paling kecil.

2.6. Model Geographically Weighted Logistic Regression (GWLR)

Geographically Weighted Logistic Regression (GWLR) adalah metode non parametrik yang merupakan bentuk lokal dari regresi logistik di mana lokasi diperhatikan dan diasumsikan bahwa data variabel dependen berdistribusi Bernoulli yang digunakan data spasial (Desriwendi, dkk., 2015). Karakteristik data spasial adalah heterogenitas spasial yang menyebabkan pengaruh variabel prediktor terhadap variabel respon pada setiap lokasi berbeda-beda, sehingga penaksiran model regresi logistik yang diaplikasikan pada data spasial yang sesuai adalah secara lokal.

Model GWLR dapat dituliskan sebagai berikut (Caraka, dkk., 2017).

$$\pi(x_i) = \frac{\exp(\sum_{j=0}^p \beta_{j(u_i, v_i)} x_{ji})}{1 + \exp(\sum_{j=0}^p \beta_{j(u_i, v_i)} x_{ji})}; i = 1, 2, \dots, n \quad (16)$$

dengan $\pi(x_i)$ adalah nilai observasi variabel dependen ke- i ; $\beta_{j(u_i, v_i)}$ adalah vektor nilai parameter variabel independent ke- j pada lokasi ke- i ; (u_i, v_i) merupakan koordinat letak geografis (*Longitude, Latitude*) dari lokasi ke- i ; dan x_{ji} adalah nilai observasi variabel independent ke- j pada lokasi ke- i .

2.7. Pemilihan Model Terbaik

Ada beberapa metode yang digunakan untuk memilih model terbaik, salah satunya *Aikake's Information Criterion* (AIC). AIC merupakan pengukuran untuk kualitas variabel dari model statistik dari data yang diberikan. Semakin kecil nilai AIC maka model akan semakin baik. Rumus AIC didefinisikan sebagai berikut (Fotheringham, dkk., 2002).

$$AIC(h) = D(h) + 2K(h) \quad (17)$$

$D(h)$ merupakan nilai devians model dengan *bandwidth* (h) dan $K(h)$ merupakan jumlah parameter dalam model dengan *bandwidth* (h). $AIC(h)$ digunakan karena kompleksitas dari model yaitu perbedaan nilai pengamatan dengan nilai dugaan, $AIC(h)$ juga tergantung pada variabel dalam model serta nilai *bandwidth*. Model terbaik adalah model dengan nilai AIC terkecil.

3. Metode Penelitian

3.1. Sumber Penelitian

Data yang digunakan pada penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik yaitu data Provinsi Sulawesi Selatan dalam Angka tahun 2022, dan data SUSENAS tahun 2022 Provinsi Sulawesi Selatan. Data pada penelitian ini dapat diakses atau diperoleh pada website <https://sulsel.bps.go.id/>.

3.2. Variabel Penelitian

Variabel respon yang digunakan adalah Angka Buta Huruf usia 15 tahun ke atas di Sulawesi Selatan tahun 2022. Variabel prediktor yang digunakan dalam penelitian ini berjumlah 4 variabel, yaitu tingkat pengangguran terbuka, tingkat kemiskinan, Angka Partisipasi Sekolah Dasar, dan persentase daerah berstatus kota.

3.3. Langkah-langkah Penelitian

1. Pendeskripsian data dengan statistika deskriptif.
2. Pendeteksian multikolinieritas variabel prediktor dengan kriteria *Variance Inflation Factor* (VIF).
3. Pengujian efek spasial.
4. Menganalisis model GWLR dengan langkah-langkah berikut.
 - a. Menghitung jarak *Euclidean* antar lokasi pengamatan berdasarkan posisi geografis (*Latitude* dan *Longitude*).
 - b. Memperoleh *bandwidth* optimum untuk setiap lokasi pengamatan dengan menggunakan *Cross Validation* (CV).
 - c. Menghitung matriks pembobot spasial menggunakan fungsi pembobot *Adaptive bisquare Kernel*, *Adaptive gaussian Kernel*, dan *Adaptive Tricube Kernel*.
 - d. Melakukan penaksiran parameter model GWLR.
5. Mendapatkan model terbaik dengan membandingkan model fungsi pembobot *Adaptive bisquare Kernel*, *Adaptive gaussian Kernel*, dan *Adaptive Tricube Kernel* pada GWLR berdasarkan nilai AIC terkecil dan menginterpretasikan model GWLR pada lokasi tertentu.
6. Menguji signifikansi parameter model GWLR untuk menentukan faktor-faktor yang berpengaruh signifikan dan interpretasi model menggunakan *odds ratio*.

4. Hasil dan Pembahasan

4.1. Deskripsi Angka Buta Huruf di Provinsi Sulawesi Selatan

Data ABH yang digunakan terdiri dari 24 Kabupaten/Kota di Provinsi Sulawesi Selatan. Dalam upaya pemilihan ABH, Kemendikbudristek menentukan target ABH 2022 adalah sebesar 1,37% penduduk. Pada penelitian ini menggunakan empat variabel prediktor yang diduga mempengaruhi ABH di Provinsi Sulawesi Selatan. Deskripsi variabel dalam statistik deskriptif yang digunakan pada penelitian ini meliputi minimum, maksimum, mean, dan standar deviasi dari setiap variabel prediktor yang dapat disajikan pada Tabel berikut.

Tabel 1 Statistik Deskriptif Variabel Prediktor

Variabel	Min.	Max.	Rerata	Simpangan Baku
TPT (X_1)	0,58	11,82	3,63	2,43
PM (X_2)	4,58	13,92	9,32	2,83
APS (X_3)	98,50	99,99	99,39	0,48
DBK (X_4)	9,09	99,35	32,03	24,72

4.2. Pendeteksian Multikolinieritas

Pendeteksian multikolinieritas dapat menggunakan kriteria nilai VIF (*Variance Inflation Factor*) yang bertujuan untuk mengetahui apakah terdapat variabel prediktor merupakan kombinasi linier atau korelasi yang tinggi dari variabel-variabel prediktor lainnya. Nilai VIF pada setiap variabel prediktor kurang dari 10 maka telah terjadi multikolinieritas antar variabel prediktor.

Tabel 2 Uji Multikolinearitas

Variabel	VIF	Indikasi Multikolinearitas
TPT (X_1)	3,8644	Tidak terdapat multikolinearitas
PM (X_2)	2,1354	Tidak terdapat multikolinearitas
APS (X_3)	1,3238	Tidak terdapat multikolinearitas
DBK (X_4)	5,0308	Tidak terdapat multikolinearitas

Berdasarkan Tabel 2 dapat dilihat bahwa hasil pengujian pada variabel TPT, PM, APS, dan DBK tidak terjadi multikolinearitas yang berarti bahwa variabel tersebut tidak saling berkorelasi.

4.3. Pengujian Efek Spasial

Pengujian efek spasial terdiri dari dua jenis yaitu ketergantungan (dependensi) spasial dan keragaman (heterogenitas) spasial.

4.3.1. Pengujian Dependensi Spasial

Pengujian dependensi spasial pada penelitian ini dilakukan untuk mengidentifikasi adanya efek spasial dependen terhadap data kasus Angka Buta Huruf di Provinsi Sulawesi Selatan dengan menggunakan uji *Moran's I* dengan hipotesis berikut.

$H_0 : \mu_i = 0$ (tidak ada dependensi antar lokasi)

$H_1 : \mu_i \neq 0$ (ada dependensi antar lokasi)

Tabel 3 Uji Moran's I

I	$E(I)$	$Var(I)$	p-value
0,21758	-0,0435	0,13789	0,05834

Berdasarkan Tabel 3, diperoleh statistik uji *Moran's I* pada pengujian dependensi spasial dimana nilai *p-value* sebesar $0,05834 > \alpha = 0,05$ maka diputuskan gagal tolak H_0 yang artinya tidak ada dependensi spasial atau pengamatan suatu lokasi tidak bergantung dengan pengamatan di lokasi lain yang letaknya berdekatan.

4.3.2. Pengujian Heterogenitas Spasial

Pengujian heterogenitas spasial bertujuan untuk mengetahui apakah data respon biner merupakan data spasial (heterogenitas spasial) dengan menggunakan uji *Breusch-Pagan*.

Tabel 4 Nilai Statistik Uji Heterogenitas Spasial

BP	$\chi^2_{(4)}$	p-value	Keputusan Uji
17,466	9,488	0,00156	Menolak H_0

Berdasarkan Tabel 4 diperoleh statistik uji *Breusch-Pagan* pada pengujian heterogenitas spasial diberikan pada persamaan (2.5) dimana nilai $BP = 17,466 > \chi^2_{(4)} = 9,488$ atau $p-value = 0,00156 < \alpha = 0,05$ maka diputuskan menolak H_0 berarti terdapat heterogenitas spasial pada data respon dikotomis.

4.4. Model Geographically Weighted Logistic Regression (GWLR)

Menghitung jarak *euclidean* antar lokasi pengamatan menggunakan titik koordinat (*Longitude* dan *Latitude*) tiap Kabupaten.Kota di Provinsi Sulawesi Selatan. Kemudian menentukan nilai *bandwidth* optimum untuk ketiga fungsi pembobot menggunakan kriteria *Cross Validation* (CV) minimum.

Tabel 5 Bandwidth Optimum dan Cross Validation

Pembobot	Bandwidth Optimum	Cross Validation
<i>Adaptive Gaussian Kernel</i>	0,3749715	1,530612
<i>Adaptive Bisquare Kernel</i>	0,0557953	1
<i>Adaptive Tricube Kernel</i>	0,2229124	1,147882

Selanjutnya, setelah menentukan *bandwidth* optimum pada masing-masing pembobot, maka dilakukan pembentukan pembobot untuk setiap lokasi sehingga didapat matriks pembobot spasial berukuran 24 x 24.

Estimasi parameter model GWLR menggunakan MLE (*Maximum Likelihood Estimation*) yang dibantu dengan metode *Newton Raphson*.

Tabel 6 Estimasi Parameter Model GWLR *Adaptive Gaussian Kernel*

	Minimum	1st Qu	Median	3st Qu	Maximum	Global
β_0	-3,8132	-1,6139	1,5674	2,4328	5,2160	4,0477
β_1	-0,0932	-0,0803	-0,0722	-0,0149	-0,0022	-0,0640
β_2	0,0004	0,0013	0,0109	0,0155	0,0215	0,0064
β_3	-0,0419	-0,0141	-0,0048	0,0275	0,0494	-0,0296
β_4	-0,0029	-0,0004	0,0001	0,0008	0,0024	0,0009

Berdasarkan ringkasan hasil estimasi parameter GWLR *Adaptive Gaussian Kernel* pada Tabel 6 memperlihatkan secara umum keragaman dari koefisien model yang dihasilkan. Untuk nilai *Intercept* atau β_0 nilai koefisien β_0 terkecil -3,8132 pada Kabupaten Barru sedangkan nilai koefisien β_0 terbesar 5,2160 pada Kabupaten Bantaeng dengan nilai mediannya 1,5674. Nilai estimasi parameter variabel lain juga memiliki bentuk interpretasi yang sama.

Tabel 7 Estimasi Parameter Model GWLR *Adaptive Bisquare Kernel*

	Minimum	1st Qu	Median	3st Qu	Maximum	Global
β_0	$-3,5785 \times 10^{+1}$	1	1	1	$1,4216 \times 10^{+1}$	4,0477
β_1	$-3,1201 \times 10^{-1}$	$-9,6645 \times 10^{-16}$	$-9,2960 \times 10^{-20}$	$1,9131 \times 10^{-17}$	$3,6968 \times 10^{-1}$	-0,0640
β_2	$-7,8466 \times 10^{-2}$	$-6,3738 \times 10^{-17}$	$-2,0692 \times 10^{-18}$	$2,6308 \times 10^{-19}$	$7,6744 \times 10^{-2}$	0,0064
β_3	$-1,174 \times 10^{-1}$	$-2,1217 \times 10^{-16}$	0	$2,1828 \times 10^{-16}$	$3,5805 \times 10^{-1}$	-0,0296
β_4	$-5,7281 \times 10^{-2}$	$-1,9553 \times 10^{-17}$	$-4,9142 \times 10^{-20}$	$1,8871 \times 10^{-16}$	$3,4757 \times 10^{-2}$	0,0009

Berdasarkan ringkasan hasil estimasi parameter GWLR *Adaptive Bisquare Kernel* pada Tabel 7 memperlihatkan secara umum keragaman dari koefisien model yang dihasilkan. Untuk nilai *Intercept* atau β_0 nilai koefisien β_0 terkecil $-3,5785 \times 10$ pada Kabupaten Barru sedangkan nilai koefisien β_0 terbesar $1,4216 \times 10$ pada Kabupaten Bantaeng dengan nilai mediannya 1. Nilai estimasi parameter variabel lain juga memiliki bentuk interpretasi yang sama.

Tabel 8 Estimasi Parameter Model GWLR *Adaptive Tricube Kernel*

	Minimum	1st Qu	Median	3st Qu	Maximum	Global
β_0	$-3,7566 \times 10$	1	1	1	$1,4029 \times 10^{+1}$	4,0477
β_1	$-3,2211 \times 10^{-1}$	$-3,7090 \times 10^{-16}$	$-2,4072 \times 10^{-17}$	$1,4115 \times 10^{-16}$	$3,7098 \times 10^{-1}$	-0,0640
β_2	$-7,8395 \times 10^{-2}$	$-3,7909 \times 10^{-16}$	$-9,4043 \times 10^{-18}$	$2,2566 \times 10^{-17}$	$8,1752 \times 10^{-2}$	0,0064
β_3	$-1,1556 \times 10^{-1}$	$-2,9597 \times 10^{-16}$	0	$3,0067 \times 10^{-16}$	$3,7506 \times 10^{-1}$	-0,0296
β_4	$-5,7425 \times 10^{-2}$	$-1,1474 \times 10^{-16}$	$-2,8047 \times 10^{-18}$	$3,1947 \times 10^{-18}$	$3,6585 \times 10^{-2}$	0,0009

Berdasarkan ringkasan hasil estimasi parameter GWLR *Adaptive Tricube Kernel* pada Tabel 8 memperlihatkan secara umum keragaman dari koefisien model yang dihasilkan. Untuk nilai *Intercept* atau β_0 nilai koefisien β_0 terkecil $-3,7566$ pada Kabupaten Barru sedangkan nilai koefisien β_0 terbesar $1,4029$ pada Kabupaten Bantaeng dengan nilai mediannya 1. Nilai estimasi parameter variabel lain juga memiliki bentuk interpretasi yang sama.

4.5. Ukuran Keباikan Model Geographically Weighted Logistic Regression (GWLR)

Perbandingan ukuran kebaikan model GWLR dengan fungsi pembobot *Adaptive Gaussian Kernel*, *Adaptive Bisquare Kernel*, dan *Adaptive Tricube Kernel* dengan menggunakan kriteria nilai AIC.

Tabel 9 Aikake's Information Criterion

Model	AIC
GWLR (Gaussian Kernel)	-41,37531
GWLR (Bisquare Kernel)	-204,3973
GWLR (Tricube Kernel)	-210.3056

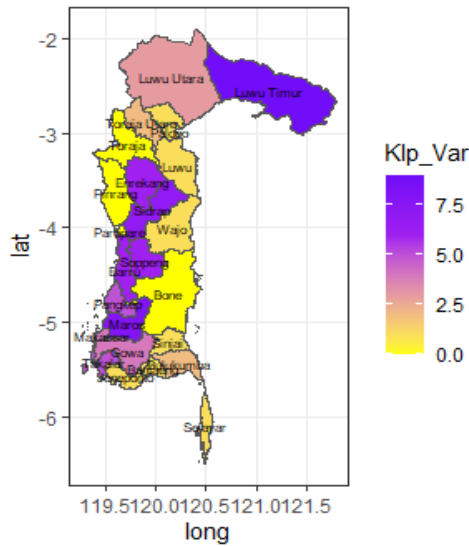
Berdasarkan Tabel 9 menunjukkan bahwa model GWLR dengan menggunakan pembobot *Adaptive Tricube Kernel* lebih baik digunakan untuk menganalisis angka buta huruf di Provinsi Sulawesi Selatan karena memiliki nilai AIC terkecil. Berdasarkan ketepatan dan nilai AIC maka model GWLR dengan pembobot *Adaptive Tricube Kernel* lebih baik digunakan untuk memodelkan angka buta huruf di Provinsi Sulawesi Selatan.

Untuk mengetahui apakah taksiran tersebut signifikan maka dilakukan perbandingan nilai $|W|$ dengan nilai $Z_{\left(\frac{\alpha}{2}\right)}$ yaitu $Z_{\left(\frac{0.05}{2}\right)} = 2,09$. Jika nilai $|W| = 2,09$ maka tolak H_0 dan variabel tersebut signifikan terhadap model. Adapun variabel yang signifikan pada setiap Kabupaten/Kota di Provinsi Sulawesi Selatan dapat dilihat pada Tabel 10 dan Gambar peta sebaran sebagai berikut.

Tabel 10 Pengelompokkan Variabel yang Signifikan

Kelompok	Kabupaten/Kota	Variabel yang Signifikan
1	Jeneponto, Luwu, Palopo, Selayar, Sinjai, Wajo, Bantaeng	$(X_1), (X_2), (X_3), \text{ dan } (X_4)$
2	Bulukumba, Toraja Utara	$(X_1), (X_2) \text{ dan } (X_3)$
3	Luwu Utara	$(X_1), (X_2) \text{ dan } (X_4)$
4	Gowa, Makassar	$(X_1) \text{ dan } (X_2)$
5	Pangkep, Takalar	$(X_1) \text{ dan } (X_3)$
6	Barru, Enrekang, Soppeng	$(X_2), (X_3), \text{ dan } (X_4)$
7	Sidrap	$(X_2) \text{ dan } (X_3)$
8	Maros	$(X_2) \text{ dan } (X_4)$
9	Luwu Timur	$(X_3) \text{ dan } (X_4)$

Berdasarkan Tabel 10, variabel prediktor yang berpengaruh terhadap angka buta huruf setiap Kabupaten/Kota di Sulawesi Selatan terbagi menjadi 9 kelompok yang berbeda-beda. Hal ini dikarenakan yang pertama, adanya heterogenitas spasial, yang berarti bahwa efek prediktor dapat bervariasi secara spasial. Dalam konteks pengelompokkan Kabupaten/Kota, ini berarti bahwa setiap wilayah dapat memiliki karakteristik unik yang menyebabkan perbedaan dalam signifikansi variabel. Kedua, memberikan estimasi parameter pada tingkat lokal, yang berarti bahwa parameter dapat berubah seiring perubahan lokasi geografis. Oleh karena itu, variabel yang signifikan pada suatu wilayah tidak signifikan pada wilayah lain. Ketiga, adanya ketergantungan spasial dalam data dapat menyebabkan variasi dalam signifikansi variabel antar wilayah. Jika suatu variabel memiliki efek yang signifikan di suatu wilayah, efek tersebut mungkin tidak sekuat atau bahkan berlawanan di wilayah lain karena adanya interaksi spasial.



Gambar 1 Pengelompokan Variabel yang Signifikan

Berdasarkan Gambar 1 menunjukkan bahwa pengelompokan yang terbentuk memiliki pola yang saling berdekatan, baik itu pada kelompok 1, 2, sampai kelompok 9.

5. Kesimpulan

Dengan melihat nilai AIC model GWLR dari ketiga fungsi pembobot, dapat dikatakan bahwa model GWLR dengan pembobot *Adaptive Tricube Kernel* mampu menggambarkan angka buta huruf Provinsi Sulawesi Selatan dengan lebih baik karena nilai AIC model GWLR dengan pembobot *Adaptive Tricube Kernel* lebih kecil. Model GWLR dengan pembobot *Adaptive Tricube Kernel* untuk Kabupaten Barru dan Kabupaten Bantaeng sebagai berikut.

$$\pi(u_1, v_1) = \frac{\exp(1,0000 + 1,6242X_{1,2}^* + 4,2047X_{1,3}^* + 2,782X_{1,4}^*)}{1 + \exp(1,0000 + 1,6242X_{1,2}^* + 4,2047X_{1,3}^* + 2,782X_{1,4}^*)}$$

$$\vdots$$

$$\pi(u_{24}, v_{24}) = \frac{\exp(1,0000 - 2,5455X_{24,1}^* + 3,4954X_{24,2}^* + 1,8258X_{24,3}^* + 2,0377X_{24,4}^*)}{1 + \exp(1,0000 - 2,5455X_{24,1}^* + 3,4954X_{24,2}^* + 1,8258X_{24,3}^* + 2,0377X_{24,4}^*)}$$

Dari hasil analisis GWLR, diperoleh faktor-faktor yang berpengaruh signifikan terhadap angka buta huruf di Provinsi Sulawesi Selatan yaitu tingkat pengangguran terbuka (X_1), penduduk miskin (X_2), angka partisipasi sekolah SD (X_3), dan daerah berstatus kota (X_4).

References

Aliu, M. A., Zubedi, F., Yahya, L., & Oroh, F. A. (2022). The Comparison of *Kernel Weighting Functions* in *Geographically Weighted Logistic Regression* in Modeling Poverty in Indonesia. *Jurnal Matematika, Statistika Dan Komputasi*, 18(3), 362–384. <https://doi.org/10.20956/j.v18i3.19567>

Amalah, R., Jaya, A. K., & Sirajang, N. (2023). Pemodelan *Geographically Weighted Logistic Regression* dengan Metode Ridge. *Journal of Statistics and Its Application*, 4(2), 130–143.

Anggraeni, Y., & Zain, I. (2015). Pemodelan Regresi Logistik Biner Terhadap Peminat ITS di Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN) 2014. *Jurnal Sains dan Seni ITS*, 4(1), 115–120.

Anisykurlillah, R. (2020). Evaluasi Pembangunan Pendidikan Keaksaraan. *Jurnal Kebijakan Pembangunan*, 15(1), 25–36.

- Anselin, L. (2003). A companion to Theoretical Econometrics Spatial econometrics. *Economics Letters*, 4, 311–329.
- Asih, I., Simbolon, R., Yatussa, F., & Wanto, A. (2021). Penerapan Algoritma Backpropagation dalam Memprediksi Persentase Penduduk Buta Huruf di Indonesia. *Jurnal Informatika UPGRIS*, 4(2), 163–169. <https://doi.org/10.26877/jiu.v4i2.2423>
- Astuti, N. K., Puhadi, P., & Andari, S. (2017). Pemodelan Angka Buta Huruf di Kabupaten/Kota se-Jawa Timur dengan Metode Geographically Weighted t Regression. *Jurnal Sains Dan Seni ITS*, 6(2), 1–53. <https://doi.org/10.12962/j23373520.v6i2.25005>
- 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. *Geographical Analysis*, 35(1), 58–82. <https://doi.org/10.1111/j.1538-4632.2003.tb01101.x>
- Badan Pusat Statistik. (2023). *Indikator Makro Sosial Ekonomi*, Sulawesi Selatan : Badan Pusat Statistik.
- Caraka, R. E., & Yasin, H. (2017). Geographically Weighted Regression (GWR). In *Encyclopedia of Geographic Information Science*, 35–70. <https://doi.org/10.4135/9781412953962.n81>
- DeMaris, A. (2005). Regression With Social Data: Modeling Continuous and Limited Response Variables. In *A John Wiley & Sons* 47(3), 282–308. <https://doi.org/10.1198/tech.2005.s304>
- Desriwendi, Abdul Hoyyi, & Triastuti Wuryandari. (2015). Pemodelan *Geographically Weighted Logistic Regression* (GWLR) dengan fungsi pembobot fixed gaussian *Kernel* dan *Adaptive gaussian Kernel* (studi kasus: laju pertumbuhan penduduk provinsi jawa tengah). *Jurnal Gaussian*, 4(2), 193–204.
- Fotheringham, Brunson, & Charlton. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationship*. Chichester: Wiley and Sons, Ltd.
- Gujarati. (2003). *Basic Econometrics, Fourth Edition*. New York : The McGraw-Hill. <https://doi.org/10.1596/1813-9450-8096>
- Hastuti, T. (2022). *Penerapan Model Geographically Weighted Logistic Regression (GWLR) Dengan Fungsi Pembobot Adaptive Gaussian Kernel Pada Data Kemiskinan Di Indonesia*. (Skripsi, Universitas Lampung). <http://digilib.unila.ac.id/68333/3>.
- Hosmer, D. ., & Lemeshow, S. (2000). *Applied Logistic Regression*. In *Biometrics* (Vol. 47, Nomor 4),1–369. <https://doi.org/10.2307/2532419>
- Leniati, A., Permana, D., Amalita, N., & Martha, Z. (2023). Analysis of the Poverty Level Model for West Sumatra Province Using Geographically Weighted Binary Logistic Regression. *Journal of Statistics and Data Science*, 1(4), 313–320.
- Lestari, V. D. (2020). *Model Geographically Weighted Logistic Regression dengan Fungsi Pembobot Adaptive Tricube*. (Skripsi, Universitas Mulawarman). <https://repository.unmul.ac.id/bitstream/handle/123456789/4833>.
- Leung, Y., Mei, C. L., & Zhang, W. X. (2000). Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A*, 32(1), 9–32. <https://doi.org/10.1068/a3162>
- Lu, B., Charlton, M., Harris, P., & Fotheringham, A. S. (2002). Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. *International Journal of Geographical Information Science*, 28(4), 660–681. <https://doi.org/10.1080/13658816.2013.865739>
- Muhyatul, H. (2016). Hakikat Pendidikan Anak usia Dini. *Jurnal Pendidikan Guru Raudhatul Athfal*, 1(1), 60-71.
- Pravitasary, Amanda, T. N. (2015). Pemodelan Faktor-Faktor yang Berpengaruh terhadap Angka Buta Huruf di Provinsi Jawa Barat dengan Geographically Weighted Logistic Regression. *Prosiding Penelitian SPeSIA*, 2(1), 21–28.
- Purnamasari, D. A., (2016). *Estimasi Parameter Model Geographically Weighted Regression (GWLR) pada Data yang Mengandung Outlier*. (Skripsi, Universitas Islam Negeri Maulana Malik Ibrahim). <http://etheses.uin-malang.ac.id/3879/1/12610010>.
- Sasmita, Winda. (2021). *Geographically Weighted Logistic Regression (GWLR) pada Pemodelan Faktor-Faktor yang Mempengaruhi Laju Pertumbuhan Penduduk Kabupaten/Kota di Provinsi Sulawesi Selatan*. (Skripsi, Universitas Negeri Makassar).
- Satvika, A. P., Luh, N., Suciptawati, P., & Susilawati, M. (2023). Memodelkan Angka Buta Huruf di Provinsi Nusa Tenggara Barat. *Journal of Education*, 06(01), 1499–1507.
- Solekha, N. A., & Qudratullah, M. F. (2022). Pemodelan *Geographically Weighted Logistic Regression* dengan Fungsi *Adaptive Gaussian Kernel* Terhadap Kemiskinan di Provinsi NTT. *Jambura Journal of Mathematics*, 4(1), 17–32.

- Sumarni, I. D., Fitriyani, N., & Baskara, Z. W. (2020). *Pemodelan Faktor yang Mempengaruhi Kemiskinan Di Provinsi Nusa Tenggara Barat Tahun 2020 Dengan Geographically Weighted Logistic Regression*. (Doctoral dissertation, Universitas Mulawarman)
- Susenas. (2022). *Statistik Pendidikan Provinsi Sulawesi Selatan*, Sulawesi Selatan: Badan Pusat Statistik.
- Tiro, M. A. (2010). *Analisis Korelasi dan Regresi Edisi Ketiga*. Makassar : Andira Publisher.
- Ulhaq, H. (2020). *Geographically Weighted Logistic Regression (Gwlr) With Gaussian Adaptive Kernel Weighting Function, Bisquare, and Tricube in Case of Malnutrition of Toddlers in Indonesia*. *Jurnal Litbang Edusaintech*, 1(1), 5–13. <https://doi.org/10.51402/jle.v1i1.2>
- UUD. (1999). *Otonomi Daerah*. Jakarta: Pemerintah Daerah.