ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 4, Nomor 2, Tahun 2015, Halaman 193 - 204
Online di: http://ejournal-s1.undip.ac.id/index.php/gaussian

PEMODELAN *GEOGRAPHICALLY WEIGHTED LOGISTIC REGRESSION* (GWLR) DENGAN FUNGSI PEMBOBOT *FIXED GAUSSIAN KERNEL* DAN *ADAPTIVE GAUSSIAN KERNEL*

(Studi Kasus : Laju Pertumbuhan Penduduk Provinsi Jawa Tengah)

Desriwendi[1], Abdul Hoyyi[2], Triastuti Wuryandari[3]
[1]Mahasiswa Jurusan Statistika FSM Undip
[2,3]Staf Pengajar Jurusan Statistika FSM Undip

ABSTRACT

The Population Growth Rate (PGR) that are not controlled will have a negative impact on the various social-economic problems such as increased poverty, crime, and so forth. Factors contributing to the population growth rate of uncontrolled allegedly various between Regency/City. Geographically Weighted Logistic Regression (GWLR) is a local form of the logistic regression where geographical factors considered. This study will analyze the factors that affect the population growth rate of Central Java Province using logistic regression and GWLR with a weighting function of Fixed Gaussian Kernel and Adaptive Gaussian Kernel. The results showed that GWLR model with a weighting function of Adaptive Gaussian Kernel better than logistic regression model and GWLR model with a weighting function of Fixed Gaussian Kernel because it has the smallest Akaike Information Criterion (AIC) value with the classification accuracy is 82.8 %.

Keywords : PGR, Logistic Regression, Fixed Gaussian Kernel, Adaptive Gaussian Kernel, GWLR, AIC.

1. PENDAHULUAN

Latar Belakang

Kepadatan penduduk adalah banyaknya penduduk terhadap satu satuan luas. Dengan mengetahui kepadatan penduduk, maka dapat diketahui konsentrasi penduduk di suatu wilayah serta dapat digunakan sebagai acuan dalam rangka mewujudkan pemerataan dan persebaran penduduk. Laju pertumbuhan penduduk Provinsi Jawa Tengah tahun 1,84 % yang dihitung berdasarkan data hasil registrasi penduduk dengan menggunakan Sistem Informasi Administrasi Kependudukan (SIAK). Nampaknya pertumbuhan ini perlu menjadi perhatian Pemerintah Provinsi maupun Pemerintah Kabupaten/Kota terkait implikasi potensi munculnya berbagai masalah sosial ekonomi seperti bertambahnya kemiskinan, kriminalitas dan lain sebagainya (BPS, 2013). Angka laju pertumbuhan penduduk dapat dikategorikan menjadi tiga yaitu apabila nilai laju pertumbuhan penduduk lebih dari nol maka terjadi pertumbuhan penduduk dari tahun sebelumnya, dan apabila nilai laju pertumbuhan sama dengan nol maka pertumbuhan penduduk tetap atau kurang dari nol maka terjadi penurunan angka pertumbuhan penduduk pada tahun tersebut dari tahun sebelumnya (www.puncakkab.bps.go.id).

Model penentuan faktor laju pertumbuhan penduduk dengan regresi logistik yang bersifat global kurang tepat diterapkan di seluruh Kabupaten/Kota di Jawa Tengah karena bisa saja suatu variabel berpengaruh terhadap pertumbuhan penduduk di satu wilayah tetapi di wilayah lain variabel tersebut tidak signifikan. Metode statistika yang digunakan untuk menganalisis heterogenitas spasial tersebut adalah *Geographically Weighted Regression* (GWR). Heterogenitas spasial adalah apabila satu variabel independen yang sama memberikan respon yang tidak sama pada lokasi yang berbeda dalam satu wilayah penelitian. Metode statistik yang juga telah dikembangkan untuk analisis data spasial yaitu *Geographically Weighted Logistic Regression* (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 non stasioner.

Tujuan Penulisan

Tujuan penelitian dari penulisan Tugas Akhir ini adalah untuk menentukan :

1. Model laju pertumbuhan penduduk Provinsi Jawa Tengah dengan menggunakan model regresi logistik.
2. Model laju pertumbuhan penduduk Provinsi Jawa Tengah dengan menggunakan GWLR dengan pembobot *Fixed Gaussian Kernel* dan *Adaptive Gaussian Kernel*.
3. Model yang mampu menggambarkan laju pertumbuhan penduduk Propinsi Jawa Tengah dengan lebih baik.

2. TINJAUAN PUSTAKA

2.1 Regresi Logistik

Model regresi logistik termasuk dalam model linier umum (*Generalized Linear Models* atau GLM). Model linier umum merupakan pengembangan dari model linier klasik. Pada model linier umum komponen acak tidak harus mengikuti distribusi normal, tetapi harus termasuk dalam distribusi keluarga eksponensial. Distribusi Binomial termasuk salah satu dalam distribusi keluarga eksponensial. Distribusi Binomial adalah distribusi dari peubah acak yang hanya mempunyai dua kategori. GLM memungkinkan pembentukan hubungan regresi antara variabel dependen dan variabel independen. Semua variabel dapat berupa variabel kontinu atau kategori, atau kombinasi dari keduanya. 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).

Hasil observasi variabel acak dependen (y) mempunyai dua kategori yaitu 0 dan 1, sehingga mengikuti distribusi Bernoulli dengan fungsi kepadatan peluang (Hosmer and Lemeshow, 1989) :

$$P(Y=y)=\pi^{y}(1-\pi)^{1-y}\ ;\ y=0,1 \qquad (1)$$

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 $\pi(x)=E(y|x)$. Model regresi logistik multivariabel dengan p variabel independen adalah :

$$\pi(x)=\frac{\exp\,(\beta_0+\beta_1x_1+\cdots+\beta_px_p)}{1+\exp\,(\beta_0+\beta_1x_1+\cdots+\beta_px_p)} \qquad (2)$$

dengan mengambil fungsi logit g(x) yaitu :

$$\mathrm{g}(x)=\ln\left[\frac{\pi(x)}{1-\pi(x)}\right]=\beta_0+\beta_1x_1+\cdots+\beta_px_p=\beta_0+\sum_{i=1}^{p}\beta_ix_i$$

Maka model regresi logistik dapat dituliskan dalam bentuk sebagai berikut :

$$\pi(\mathrm{x}) = \frac{\exp(\mathrm{g}(\mathrm{x}))}{1+\exp(\mathrm{g}(\mathrm{x}))} \qquad (3)$$

2.2 Penaksir Parameter Model Regresi Logistik

Untuk menentukan estimasi parameter $\boldsymbol{\beta}$ digunakan metode *Maximum Likelihood Estimation* (MLE) dengan iterasi Newton Raphson yang merupakan penyelesaian dari turunan pertama dari fungsi likelihood. Penaksiran varian dan kovarian diperoleh dari turunan kedua fungsi ln likelihood. Misalnya Y_i adalah variabel acak berdistribusi binomial dengan $E(Y_i) = n_i\pi(x_i)$ dan $i = 1, \dots, I$, dimana $n_1 + \dots + n_I$ dengan fungsi densitas probabilitas bersama (Agresti, 1990) adalah : $L(\beta) = \prod_{i=1}^{I} P(Y = y_i) = \prod_{i=1}^{I} \pi(x_i)^{y_i}(1-\pi(x_i))^{n_i-y_i}$

$$= \{ \prod_{i=1}^{I}(1-\pi(x_i))^{n_i}\}\exp\left[\sum_{i=1}^{I} y_i \ln\left(\frac{\pi(x_i)}{(1-\pi(x_i))}\right)\right]$$

Untuk memudahkan perhitungan, maka fungsi likelihood dimaksimumkan dalam bentuk ln $L(\beta)$.

$$\ln L(\boldsymbol{\beta}) = \sum_{k=0}^{p}(\sum_{i=1}^{I} y_i x_{ik})\beta_k - \sum_{i=1}^{I} n_i \ln\{1 + \exp\,(\sum_{k=0}^{p} \beta_k x_{ik})\}$$

Menurunkan ln $L\,(\boldsymbol{\beta})$ terhadap β_k dan hasilnya sama dengan nol

$$\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_k} = \sum_{i=1}^{I} y_i x_{ik} - \sum_{i=1}^{I} n_i x_{ik}\pi(x_i) = 0$$

Penaksiran varian dan kovarian diperoleh dari turunan kedua dari fungsi likelihood (Agresti, 1990) :

$$\frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_k \partial \beta_{k^*}} = -\sum_{i=1}^{I} x_{ik}x_{ik^*}\pi(x_i)(1-\pi(x_i))\,; \qquad k, k^* = 1,2, \dots, p$$

Nilai parameter $\boldsymbol{\beta}$ dari turunan pertama fungsi $L(\boldsymbol{\beta})$ diperoleh melalui suatu prosedur iterative yang dilakukan dengan metode iterasi Newton Raphson yaitu memaksimumkan fungsi likelihood (Agresti, 1990).

2.3 Pengujian Parameter Model Regresi Logistik

a. Uji Secara Simultan

Menurut Hosmer dan Lemeshow (1989) untuk menguji signifikansi parameter β dalam model secara bersama-sama dengan menggunakan statistik uji *G*.

Hipotesis :

$\mathrm{H}_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$

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

Taraf signifikansi : α

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

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 $G > \chi^2_{(a,p)}$ dengan *p* adalah banyaknya variabel independen

b. Uji Parsial

Menurut Hosmer dan Lemeshow (1989) untuk mengetahui signifikansi parameter β terhadap variabel dependennya secara parsial menggunakan statistik uji Wald.

Hipotesis :

$$\mathrm{H}_0 : \beta_j = 0 \; ; \quad j = 1{,}2, \dots , p$$
$$\mathrm{H}_1 : \beta_j \neq 0 \; ; \quad j = 1, 2, \dots , p$$

Taraf signifikansi : α

Statistik uji : $Z_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$ atau $W_j^2 = \frac{\hat{\beta}_j^{\,2}}{(SE(\hat{\beta}_j))^2}$ 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 $W_j^2 > \chi^2{}_{(a,1)}$

2.4 Diagnostik Multikolinieritas

Menurut Sarwoko (2005) ada beberapa metode untuk mendeteksi multikolinieritas diantaranya :

1. Kolinieritas seringkali diduga ketika R^2 tinggi (misalnya : antara 0,7 dan 1) dan jika R^2 tinggi, ini akan 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.
2. Multikolinieritas timbul karena satu atau lebih variabel yang menjelaskan merupakan kombinasi liner yang mendekati pasti dari variabel yang menjelaskan lainnya.

2.5 Model *Geographically Weighted Regression* (GWR)

Menurut Fotheringham, *et al.* (2002) GWR adalah metode statistika yang digunakan untuk menganalisis heterogenitas spasial. Model GWR dapat ditulis sebagai berikut :

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \; ; \; i = 1, 2, \dots , n \tag{4}$$

2.5.1 Pemilihan Pembobot (*Weight*)

a. Fungsi invers jarak

Misalkan $1/d_{ij}$ adalah fungsi invers jarak yang mewakili pembobot antara lokasi ke-*i* dan lokasi ke-*j* dimana $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ adalah jarak *euclidean* antara lokasi *i* dan lokasi *j* (Anggarini dan Purhadi, 2012). 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* (Fotheringham, *et al.*, 2002).

Pembobot ini dapat ditulis:

$$w_j(u_i, v_i) = \begin{cases} 1, \text{jika } d_{ij} \leq r \\ 0, \text{jika } d_{ij} > r \end{cases} \tag{5}$$

b. Fungsi Kernel

Fungsi pembobotnya masing-masing dapat ditulis sebagai berikut :

1. *Fixed Gaussian Kernel* (Fotheringham, *et al.*, 2002) :

$w_{ij} = \exp[-((d_{ij}/h)/2)^2]$ (6)

Dimana *h* merupakan *bandwidth* yang *fixed* atau *bandwidth* yang yang sama digunakan untuk setiap lokasi.

2. *Adaptive Gaussian Kernel* (Anggarini dan Purhadi, 2012) :

$w_{ij} = \exp[-((d_{ij}/h_{i(q)})/2)^2]$ (7)

dengan *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*.

2.5.2 Pemilihan *Bandwidth* Optimum dan Model Terbaik

Menurut Fotheringham, *et al.* (2002) pemilihan ukuran *bandwidth* yang optimum menjadi salah satu hal yang penting karena akan mempengaruhi ketepatan hasil regresi. Pada penelitian ini digunakan metode *Cross Validation* (CV) dengan rumus sebagai berikut :

$$CV = \sum_{i=1}^{n}[y_i - \hat{y}_{\neq i}(h)]^2 \quad (8)$$

Dengan $\hat{y}_{\neq i}(h)$ adalah nilai penaksir y_i dimana pengamatan di lokasi i dihilangkan dari proses penaksiran. Untuk pemilihan model terbaik ditentukan dengan nilai *Akaike Information Criterion* (AIC) yang paling kecil dengan persamaan berikut :

$$AIC = 2n\,log_e(\hat{\sigma}) + n\,log_e(2\pi) + n + tr(\boldsymbol{S}) \quad (9)$$

2.6 Penaksiran Parameter $\boldsymbol{\beta}(u_i, v_i)$

Menurut Fotheringham, *et al.* (2002), metode penaksiran parameter pada model GWR adalah dengan metode *Weighted Least Square* (WLS) yaitu dengan memberikan pembobot yang berbeda untuk setiap lokasi dimana data tersebut dikumpulkan. Misalkan pembobot untuk setiap lokasi ke-i adalah $w_j(u_i, v_i)$; *j*=1,2,…,*n*, maka parameter lokasi (u_i, v_i) ditaksir dengan menambahkan unsur pembobot pada persamaan (4) dan kemudian meminimumkan jumlah kuadrat error berikut ini :

$$\sum_{j=1}^{n} w_j(u_i,v_i)\varepsilon_j^2 = \sum_{j=1}^{n} w_j(u_i,v_i)(y_j - \beta_0(u_i,v_i) - \beta_1(u_i,v_i)x_{j1} - \beta_2(u_i,v_i)x_{j2} - \ldots - \beta_p(u_i,v_i)x_{jp})^2$$

Bentuk penaksir parameter dari model GWR untuk setiap lokasi adalah :

$$\hat{\boldsymbol{\beta}}(u_i,v_i) = \left(\mathbf{X}^T\mathbf{W}(u_i,v_i)\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{W}(u_i,v_i)\mathbf{y} \quad (10)$$

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 GWR (*Geographically Weighted Regression*) yang menggabungkan parameter non stasioner dan data kategorikal. Dalam penelitian ini, GWR dan regresi logistik digabungkan untuk membentuk GWLR. 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:

$$\pi(x_i) = \frac{\exp\left(\sum_{k=0}^{p}\beta_k(u_i,v_i)x_{ik}\right)}{1+\exp\left(\sum_{k=0}^{p}\beta_k(u_i,v_i)x_{ik}\right)} \;;\; i = 1,2,\ldots,n \text{ dan } x_{i0} = 1 \quad (11)$$

Penaksiran parameter model GWLR dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Penaksiran parameter diperoleh dengan memaksimumkan fungsi logaritma natural likelihood dengan cara menurunkannya terhadap $\beta_k(u_i, v_i)$ dan hasilnya disamakan dengan nol. Langkah pertama yang harus dilakukan yaitu membentuk fungsi ln likelihoodnya sebagai berikut :

$$\ln L\big(\beta(u_i, v_i)\big) = \sum_{k=0}^{p}\left(\sum_{i=1}^{I} y_i x_{ik}\right)\beta_k(u_i, v_i) - \sum_{i=1}^{I} n_i \ln\left\{1 + \exp\left(\sum_{k=0}^{p}\beta_k(u_i, v_i)x_{ik}\right)\right\}$$

pembobot diberikan pada fungsi logaritma natural likelihood untuk mendapatkan model GWLR.

$$\ln L^*\big(\beta(u_i, v_i)\big) = \sum_{k=0}^{p}\left(\sum_{i=1}^{I} w_i(u_i, v_i) y_i x_{ik}\right)\beta_k(u_i, v_i)$$
$$- \sum_{i=1}^{I} n_i w_i(u_i, v_i) \ln\left\{1 + \exp\left(\sum_{k=0}^{p}\beta_k(u_i, v_i)x_{ik}\right)\right\}$$

persamaan fungsi logaritma natural likelihood diturunkan terhadap $\beta_k(u_i, v_i)$ dan disamakan dengan nol $\frac{\partial \ln L^*(\beta_k(u_i,v_i))}{\partial(\beta_k(u_i,v_i))} = 0$ yang dapat diselesaikan dengan metode *Iteratively Reweighted Least Square* (IRLS) (Anggarini dan Purhadi, 2012).

2.8 Pengujian Kesesuaian Model Regresi Logistik dan Model GWLR

Pengujian kesamaan model Regresi Logistik dan model GWLR menggunakan perbandingan nilai *deviance* model Regresi Logistik dan model GWLR dengan hipotesis sebagai berikut :

H_0 : $\beta_k(u_i, v_i) = \beta_k$, $i = 1,2,\ldots,n$, $k = 1,2,\ldots,p$ (Tidak ada perbedaan yang signifikan antara model Regresi Logistik dan model GWLR)

H_1 : paling sedikit satu $\beta_k(u_i, v_i) \neq \beta_k$, $i = 1,2,\ldots,n$, $k = 1,2,\ldots,p$ (Ada perbedaan yang signifikan antara model Regresi Logistik dan model GWLR)

Pengujian kesamaan model Regresi Logistik dan model GWLR menggunakan perbandingan nilai *deviance* model Regresi Logistik dan model GWLR. Misalkan model Regresi Logistik *deviance* dinyatakan dengan $D(\hat{\beta})$ dengan derajat bebas v_1 dan model GWLR dinyatakan dengan $D(\hat{\beta}(u_i, v_i))$ dengan derajat bebas v_2 maka :

Statistik Uji :

$$F_{hit} = \frac{D(\hat{\beta})/v_1}{D(\hat{\beta}(u_i, v_i))/v_2}$$

Deviance Regresi Logistik dirumuskan (Atkinson, *et al*, 2003) :

$$D\big(\hat{\beta}\big) = -2\sum_{i=1}^{n}\pi(x_i)logit\big(\pi(x_i)\big) + \log\,(1 - \pi(x_i))$$

Deviance GWLR dirumuskan :

$$\mathrm{D}(\hat{\beta}(u_i, v_i)) = 2\,\{\textstyle\sum_{i=1}^{n}[y_{1i}\ln\hat{\pi}(x_i) + y_{0i}\ln(1 - \hat{\pi}(x_i))]\} -$$
$$2\{\textstyle\sum_{i=1}^{n} n_{1i}\ln(n_{1i}) + n_{0i}\ln(n_{0i}) + n\ln(n)\}$$

Akan mengikuti distribusi F dengan derajat bebas v_1 dan v_2

Kriteria Uji : tolak H_0 jika $F_{hit} > F_{(\alpha, v_1, v_2)}$

2.9 Pengujian Parameter Model GWLR

Untuk menguji parameter $\beta(u_i, v_i)$ secara parsial yang bertujuan untuk mengetahui parameter mana saja yang signifikan mempengaruhi variabel dependennya (Lailiyah dan Purhadi, 2012).

Hipotesis :

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

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

Statistik uji :

$$Z_{hit} = \frac{\hat{\beta}_k(u_i,v_i)}{SE(\hat{\beta}_k(u_i,v_i))} ;$$

dimana $SE\left(\hat{\beta}_k(u_i, v_i)\right) = \sqrt{var(\hat{\beta}_k(u_i, v_i))}$

Kriteria Uji : Tolak H_0 jika $|Z_{hit}| > Z_{\alpha/2}$

2.10 Laju Pertumbuhan Penduduk

2.10.1 Definisi

Laju pertumbuhan penduduk adalah rata-rata tahunan laju perubahan jumlah penduduk di suatu daerah selama periode waktu tertentu sehingga dapat diketahui perubahan jumlah penduduk antar dua periode waktu. (BPS, 2013).

2.10.2 Faktor-Faktor yang Mempengaruhi Pertumbuhan Penduduk

Banyak faktor yang menentukan besarnya penduduk suatu kota serta laju pertumbuhannya dari waktu ke waktu yang satu sama lainnya saling terkait dan saling mempengaruhi. Kalau dilihat dari sudut demografi, faktor tersebut terdiri dari tiga unsur yaitu fertilitas (kelahiran), mortalitas (kematian), dan migrasi (perpindahan) yang masing-masing dipengaruhi berbagai faktor (BPS, 2000).

2.10.3 Menghitung Laju Pertumbuhan Penduduk

Laju pertumbuhan penduduk digunakan untuk mengetahui perubahan jumlah penduduk antar dua periode waktu. Berikut rumus perhitungan laju pertumbuhan penduduk (www.puncakkab.bps.go.id) dengan metode geometri :

$$r = \left(\frac{P_t}{P_0}\right)^{1/t} - 1 \qquad (12)$$

Kategori laju pertumbuhan penduduk :

1. Laju pertumbuhan penduduk > 0 berarti terjadi penambahan jumlah penduduk
2. Laju pertumbuhan penduduk = 0 berarti tidak ada perubahan jumlah penduduk
3. Laju pertumbuhan penduduk < 0 berarti terjadi pengurangan jumlah penduduk.

3. METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) dalam buku Jawa Tengah Dalam Angka tahun 2013 dan 2014

3.2 Variabel Penelitian

Variabel dependen (Y) pada penelitian ini adalah angka Laju Pertumbuhan Penduduk (LPP) tiap Kabupaten atau Kota di Jawa Tengah. Sedangkan variabel independennya (X) yaitu jumlah kelahiran (X1), jumlah kematian (X2), dan Jumlah

Migran (X3) dengan kategori Migran masuk = 1 dan Migran keluar = 0 tiap Kabupaten/Kota di Jawa Tengah tahun 2013.

3.3 Metode Penelitian

Berikut langkah-langkah analisis data dalam penelitian ini :

1. Mengambil data jumlah penduduk tahun 2012 dan 2013 sebagai variabel dependen (Y) dan keempat variabel independennya serta data geografis letak titik koordinat masing-masing Kabupaten/Kota di Jawa Tengah.
2. Menghitung laju pertumbuhan penduduk tiap-tiap Kabupaten/Kota untuk mengkategorikan ke dalam kategori terjadi penambahan penduduk, jumlah penduduk tetap, atau terjadi penurunan penduduk.
3. Memeriksa kolinieritas dari variabel-variabel independen
4. Membuat model regresi logistik dan menguji signifikansi parameter
5. Menganalisis model GWLR dengan langkah-langkah sebagai berikut:
   a. Menentukan u_i dan v_i berdasarkan garis lintang selatan dan garis bujur timur untuk masing-masing Kabupaten/Kota di Jawa Tengah.
   b. Menentukan nilai *bandwidth* menggunakan metode *Cross Validation* (CV)
   c. Menghitung jarak *Euclidian* antar lokasi pengamatan berdasarkan letak geografis.
   d. Menghitung matriks pembobot menggunakan fungsi Kernel.
   e. Mendapatkan penaksir parameter model GWLR.
   f. Melakukan uji signifikansi parameter.
6. Melakukan uji kesesuaian model regresi logistik dan GWLR.
7. Mendapatkan model terbaik yang mempunyai nilai AIC terkecil dengan membandingkan nilai AIC model regresi logistik dengan GWLR .
8. Membuat kesimpulan.

4. HASIL DAN PEMBAHASAN

4.1 Deskripsi Data Penelitian

Berikut deskripsi hasil data variabel-variabel yang mempengaruhi Laju Pertumbuhan Penduduk Provinsi Jawa Tengah tahun 2013 :

Tabel 1 Deskripsi Data Penelitian

Variabel	Rata-rata	Simpangan baku	Min	Maks
X1	16089.57	7407.989	1798	33074
X2	13116.97	7052.844	1704	31428
X3	0.2	0.4	0	1

4.2 Model Regresi Logistik

model awal regresi logistik untuk laju pertumbuhan penduduk Provinsi Jawa Tengah tahun 2013 sebagai berikut :

$$\hat{\pi}(\mathrm{x}) = \frac{\exp(-0{,}21407+ 0{,}0003395\mathrm{X}_1 - 0{,}0004518\mathrm{X}_2 + 2{,}46245\mathrm{X}_3(1)\ )}{1 + \exp(-0{,}21407+ 0{,}0003395\mathrm{X}_1 - 0{,}0004518\mathrm{X}_2 + 2{,}46245\mathrm{X}_3(1)\ )}$$

parameter yang berpengaruh secara signifikan pada $\alpha = 5\ \%$ yaitu pada variabel X_1 dan X_2 (Jumlah Kelahiran dan Kematian tiap Kabupaten/Kota di Jawa Tengah tahun 2013 karena nilai |Z| = 2,13 > $Z_{(0.025)} = 1{,}96$ dan |Z| = 2,38 > $Z_{(0.025)} = 1{,}96$

4.3 Model *Geographically Weighted Logistic Regression* (GWLR)

Langkah pertama dalam mendapatkan model GWLR adalah menentukan letak geografis (garis lintang dan garis bujur) tiap Kabupaten/Kota di Jawa Tengah. Selanjutnya adalah menghitung *bandwidth* optimum dengan menggunakan metode *Cross Validation* (CV). Proses untuk mendapatkan *bandwidth* yang meminimumkan nilai CV bisa dilakukan dengan menggunakan teknik *Golden Section Search* (Fotheringham, *et al.*, 2002). Setelah mendapatkan nilai *bandwidth* optimum, maka langkah selanjutnya adalah menentukan matriks pembobot, dimana dalam penelitian ini akan digunakan dua pembobot yaitu fungsi *Fixed Gaussian Kernel* dan fungsi *Adaptive Gaussian Kernel*. Misalkan matriks pembobot untuk Kabupaten Cilacap di lokasi (u_i, v_i) adalah $w(u_i, v_i)$. Langkah awal sebelum mendapat matriks pembobot ini adalah dengan mencari jarak *Euclidian* lokasi (u_i, v_i).

Misalkan matriks pembobot yang digunakan untuk menaksir parameter hanya pada lokasi (u_1, v_1) yaitu Kabupaten Cilacap, sedangkan untuk menaksir parameter pada lokasi lainnya yaitu lokasi (u_2, v_2) perlu dicari terlebih dahulu matriks pembobot $w(u_2, v_2)$ dengan langkah yang sama seperti cara sebelumnya, demikian seterusnya sampai pada lokasi (u_{35}, v_{35}). Pembobot tersebut akan digunakan untuk mencari penaksir parameter model GWLR dengan memasukkan pembobot tersebut dalam perhitungannya.

Setelah diperoleh penaksiran parameter model regresi logistik dan model GWLR, maka langkah selanjutnya adalah mengetahui ada tidaknya perbedaan yang signifikan antara model regresi logistik dengan model GWLR sebagai berikut :

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

H_1 : Paling sedikit satu $\beta_k(u_i, v_i) \neq \beta_k$; $k = 1, 2, 3$; $i = 1, 2, \dots, 35$

Pengujian kesamaan model dilakukan dengan menggunakan uji F dan diperoleh hasil sebagai berikut:

Tabel 2 Uji Kesesuaian Model Regresi Logistik dan Model GWLR

Model	*Deviance*	DOF	*Deviance*/DOF	F_{hit}
Regresi Logistik	35.307	31.000	1.139	
GWLR (Fixed Gaussian)	29.144	27.628	1.055	1.079
GWLR (Adaptif Gaussian)	31.530	29.411	1.072	1.063

Tabel 2 menunjukkan nilai F_{hit} dengan menggunakan pembobot *Fixed Gaussian Kernel* dan *Adaptive Gaussian Kernel* masing-masing adalah 1,079 dan 1,063. Apabila menggunakan α = 0,05 maka nilai $F_{0,05;31;27.628} = 1{,}878$ dan $F_{0,05;31;29,411} = 1{,}848$. Dari hasil tersebut maka dapat disimpulkan bahwa H_0 diterima yang artinya tidak ada perbedaan yang signifikan antara model regresi logistik dengan model GWLR dengan kedua pembobot. Dalam pemilihan model terbaik dapat dilihat dari nilai AIC. Model yang mempunyai AIC paling kecil tersebut yang merupakan model terbaik untuk Laju Pertumbuhan Penduduk Provinsi Jawa Tengah tahun 2013.

4.3.1 Pengujian Parameter Model GWLR Pembobot *Fixed Gaussian Kernel*

Pengujian parameter model GWLR dengan pembobot *Fixed Gaussian Kernel* digunakan untuk mengetahui faktor-faktor yang berpengaruh terhadap Laju Pertumbuhan Penduduk Provinsi Jawa Tengah tahun 2013 di setiap Kabupaten/Kota.

Misalkan pengujian parameter untuk Kabupaten Cilacap yang lokasinya pada koordinat (u_1, v_1) maka hipotesisnya adalah sebagai berikut:

$\mathrm{H}_0 : \beta_k(u_1, v_1) = 0;$
$\mathrm{H}_1 : \beta_k(u_1, v_1) \neq 0;$ $k = 1, 2, 3$

Tabel 3 Pengujian Parameter Model GWLR Kabupaten Cilacap dengan Pembobot *Fixed Gaussian Kernel*

Parameter	Estimasi	*Standard Error*	Z_{hit}
β_0	-0.27301	0.80066	-0.34098
β_1	0.00030	0.00009	3.39635
β_2	-0.00040	0.00009	-4.51434
β_3	2.41227	0.67938	3.55070

Berdasarkan Tabel 3 maka diperoleh model GWLR dengan pembobot *Fixed Gaussian Kernel* untuk Kabupaten Cilacap yaitu :

$$\hat{\pi}(\mathrm{x}) = \frac{\exp(-0{,}27301 + 0{,}00030\mathrm{X}_1 - 0{,}00040\mathrm{X}_2 + 2{,}41227\mathrm{X}_3)}{1 + \exp(-0{,}27301 + 0{,}00030\mathrm{X}_1 - 0{,}00040\mathrm{X}_2 + 2{,}41227\mathrm{X}_3)}$$

Proses pengujian parameter tersebut dilakukan berulang pada setiap lokasi yaitu sampai lokasi (u_{35}, v_{35}) atau sampai Kota Tegal. Variabel yang berpengaruh secara signifikan yaitu variabel yang mempunyai nilai $|Z_{\text{hit}}| > Z_{(0.025)} = 1{,}96$. Variabel yang berpengaruh secara signifikan dalam model GWLR dengan pembobot *Fixed Gaussian Kernel* yaitu jumlah kelahiran (X1), jumlah kematian (X2), dan jumlah Migran (X3) tiap Kabupaten/Kota di Provinsi Jawa Tengah tahun 2013

4.3.2 Pengujian Parameter Model GWLR dengan *Adaptive Gaussian Kernel*

Pengujian parameter model GWLR dengan pembobot *Adaptive Gaussian Kernel* digunakan untuk mengetahui faktor-faktor yang berpengaruh terhadap Laju Pertumbuhan Penduduk Provinsi Jawa Tengah tahun 2013 di setiap Kabupaten/Kota. Misalkan pengujian parameter untuk Kabupaten Cilacap yang lokasinya pada koordinat (u_1, v_1) maka hipotesisnya adalah sebagai berikut:

$\mathrm{H}_0 : \beta_k(u_1, v_1) = 0;$
$\mathrm{H}_1 : \beta_k(u_1, v_1) \neq 0;$ $k = 1, 2, 3$

Tabel 4 Pengujian Parameter Model GWLR Kabupaten Cilacap dengan Pembobot *Adaptive Gaussian Kernel*

Parameter	Estimasi	*Standard Error*	Z_{hit}
β_0	-0.28403	0.75293	-0.37724
β_1	0.00029	0.00009	3.34508
β_2	-0.00039	0.00008	-4.75271
β_3	2.40135	0.66404	3.61627

Berdasarkan Tabel 4 diperoleh model GWLR dengan pembobot *Adaptive Gaussian Kernel* untuk Kabupaten Cilacap yaitu :

$$\hat{\pi}(\mathrm{x}) = \frac{\exp(\text{-}0{,}28403 + 0{,}00029\mathrm{X}_1 - 0{,}00039\mathrm{X}_2 + 2{,}40135\mathrm{X}_3)}{1 + \exp(\text{-}0{,}28403 + 0{,}00029\mathrm{X}_1 - 0{,}00039\mathrm{X}_2 + 2{,}40135\mathrm{X}_3)}$$

Proses pengujian parameter tersebut dilakukan berulang pada setiap lokasi yaitu sampai lokasi (u_{35}, v_{35}) atau sampai Kota Tegal. Variabel yang berpengaruh secara signifikan yaitu variabel yang mempunyai nilai $|Z_{\text{hit}}| > Z_{(0.025)} = 1{,}96$. Variabel yang berpengaruh secara signifikan dalam model GWLR dengan pembobot *Adaptive Gaussian Kernel* yaitu jumlah kelahiran (X1), jumlah kematian (X2), dan jumlah Migran (X3) tiap Kabupaten/Kota di Provinsi Jawa Tengah tahun 2013.

4.3.3 Perbandingan Model Regresi Logistik dan Model GWLR

Perbandingan antara model regresi logistik dan model GWLR dengan kedua pembobot *Fixed Gaussian Kernel* dan *Adaptive Gaussian Kernel* untuk mengetahui model yang lebih baik dalam menggambarkan Laju Pertumbuhan Penduduk di Provinsi Jawa Tengah tahun 2013. Perbandingan ini dapat dilihat dari besarnya nilai AIC dari masing-masing model tersebut yaitu :

Tabel 5 Perbandingan Kesesuaian Model

Model	*Deviance*	AIC
Regresi Logistik	35,307	43,307
GWLR (*Fixed Gaussian Kernel*)	29,144	42,500
GWLR (*Adaptive Bisquare Kernel*)	31,530	41,541

Dari hasil analisis pada Tabel 5 dapat dilihat bahwa nilai AIC terkecil dimiliki oleh model GWLR dengan pembobot *Adaptive Gaussian Kernel* yang artinya model GWLR dengan pembobot *Adaptive Gaussian Kernel* adalah model yang lebih baik daripada model regresi logistik dan model GWLR dengan pembobot *Fixed Gaussian Kernel*.

4.3.4 Ketepatan Klasifikasi Model Regresi Logistik dan GWLR

Perhitungan ketepatan hasil klasifikasi Laju Pertumbuhan Penduduk Provinsi Jawa Tengah Tahun 2013 dengan menggunakan model regresi logistik, GWLR pembobot *Fixed Gaussian Kernel* dan GWLR pembobot *Adaptive Gaussian Kernel* menghasilkan nilai yaitu sebesar 77,1%, 80,0% dan 82.8%. Berdasarkan nilai ketepatan hasil klasifikasi tersebut, model GWLR untuk pembobot *Adaptive Gaussian Kernel* merupakan model yang terbaik karena memiliki nilai *Akaike Information Criterion* (AIC) yang terkecil dengan ketepatan model sebesar 82.8 %.

5. KESIMPULAN

Kesimpulan yang didapat dari hasil dan pembahasan pada penelitian ini sebagai berikut :

1. Model regresi logistik laju pertumbuhan penduduk Provinsi Jawa Tengah tahun 2013 yaitu:
$$\hat{\pi}(\mathrm{x}) = \frac{\exp(\text{-}0{,}21407 + 0{,}0003395\mathrm{X}_1 - 0{,}0004518\mathrm{X}_2 + 2{,}46245\mathrm{X}_3(1))}{1 + \exp(\text{-}0{,}21407 + 0{,}0003395\mathrm{X}_1 - 0{,}0004518\mathrm{X}_2 + 2{,}46245\mathrm{X}_3(1))}$$
Variabel independen yang mempengaruhi secara signifikan terhadap laju pertumbuhan penduduk Provinsi Jawa Tengah tahun 2013 yaitu variabel jumlah kelahiran dan kematian pada tiap Kabupaten/Kota di Jawa Tengah.
2. Model GWLR dengan pembobot *Fixed Gaussian Kernel* dan pembobot *Adaptive Gaussian Kernel* variabel yang berpengaruh secara signifikan terhadap laju pertumbuhan penduduk tiap Kabupaten/Kota Provinsi Jawa Tengah tahun 2013

yaitu variabel jumlah kelahiran, jumlah kematian, dan jumlah Migran untuk setiap lokasi.

3. Model GWLR dengan pembobot *Adaptive Gaussian Kernel* merupakan model yang lebih baik untuk menggambarkan laju pertumbuhan penduduk Provinsi Jawa Tengah tahun 2013 karena memiliki nilai AIC yang lebih kecil yaitu 41,541 dengan ketepatan model sebesar 82,8 %.

6. DAFTAR PUSTAKA

Agresti, A. 1990. *Categorical Data Analysis, Second Edition.* John Wiley & Sons : New York.

Agresti, A. 2002. *Categorical Data Analysis, Second Edition.* John Wiley & Sons : New York

Atkinson, P. M., S. E. German, D. A. Sear, and M. J. Clark. 2003. *Exploring The Relations Between Riverbank Erosion and Geomorphological Controls Using Geographically Weighted Logistic Regression.* Ohio State University, Ohio

Fotheringham, A.S. Brundson, C. dan Charlton, M. 2002. *Geographically Weighted Regression : Analysis of Spatially Varying Relationship.* John Wiley and Sons Ltd : England.

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

Prasetyo, E. 2012. *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta : ANDI

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

Wirosuhardjo, K. 1981. *Dasar-Dasar Demografi*. FEUI : Jakarta.