

Pemodelan Faktor-Faktor yang Mempengaruhi Jumlah Kasus Diabetes Melitus di Jawa Timur Menggunakan GWGPR dan GWNBR

Elvira Dian Safire^{1*} dan Puhadi²

^{1,2}Departemen Statistika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember

*Corresponding author: elviradiansafire@gmail.com

Received: 31 March 2022

Revised: 24 September 2023

Accepted: 20 March 2023

ABSTRAK – Diabetes melitus merupakan penyakit kronis berupa gangguan metabolik yang ditandai dengan kadar gula darah yang melebihi batas normal. Provinsi penyumbang jumlah kasus penderita diabetes melitus terbanyak di Indonesia pada tahun 2019 adalah Provinsi Jawa Timur. Untuk mengetahui faktor-faktor yang mempengaruhi jumlah kasus diabetes melitus digunakan pendekatan dengan metode *Geographically Weighted Generalized Poisson Regression* (GWGPR) dan *Geographically Weighted Negative Binomial Regression* (GWNBR). Jumlah penderita diabetes melitus terbanyak di Jawa Timur adalah di Kota Surabaya dengan jumlah 94.076 kasus dan terendah ada di Kota Batu yakni 3.344 kasus. Pemodelan GWGPR menghasilkan 3 kelompok dan GWNBR menghasilkan 4 kelompok untuk variabel yang signifikan pada masing-masing kabupaten/kota. Faktor-faktor yang berpengaruh signifikan terhadap jumlah kasus diabetes melitus antara lain rasio pendidikan SMA ke-atas, proporsi aktivitas fisik penduduk, persentase konsumsi rokok tembakau, persentase penderita hipertensi, dan proporsi konsumsi minuman manis. Model GWGPR memiliki nilai *Akaike's Information Criteria Corrected* (AICc) yang lebih kecil daripada model GWNBR, sehingga metode GWGPR paling sesuai untuk memodelkan jumlah kasus diabetes melitus di kabupaten/kota di Jawa Timur dibandingkan metode GWNBR.

Keywords – Diabetes Melitus, GWGPR, GWNBR

ABSTRACT – Diabetes mellitus is a chronic disease of metabolic disorders characterized by blood-sugar levels exceeding normal limits. The province contributing the largest number of cases of diabetes mellitus in Indonesia in 2019 is East Java Province. To know the factors that affect the number of cases of diabetes mellitus is used approach with *Geographically Weighted Generalized Poisson Regression* (GWGPR) and *Geographically Weighted Negative Binomial Regression* (GWNBR) methods. The highest number of people with diabetes mellitus in East Java is in Surabaya with 94076 cases and the lowest is in Batu City which is 3344 cases. GWGPR and GWNBR modeling both resulted in 4 groups for significant variables in each district / city. The AICc value comparison of the GWGPR and GWNBR models shows almost the same value. Sehingga menunjukkan bahwa model GWGPR dan GWNBR sudah sesuai. So, it shows that the GWGPR and GWNBR models are appropriate. The GWGPR model has a smaller AICc value than the GWNBR model, so the GWGPR method is best suited to model the number of cases of diabetes mellitus in districts/cities in East Java compared to GWNBR method.

Keywords – Diabetes Melitus, GWGPR, GWNBR

I. INTRODUCTION

Diabetes merupakan penyakit kronis berupa gangguan metabolik yang ditandai dengan kadar gula darah yang melebihi batas normal. Organisasi Internasional Diabetes (IDF) memperkirakan setidaknya terdapat 463 juta orang pada usia 20-79 tahun di dunia menderita diabetes pada tahun 2019 atau setara dengan angka prevalensi sebesar 9,3% dari total penduduk pada usia yang sama. Prevalensi diabetes melitus menurut hasil pemeriksaan gula darah meningkat dari 6,9% pada 2013 menjadi 8,5% pada tahun 2018. Angka ini menunjukkan bahwa baru sekitar 25% penderita diabetes yang mengetahui bahwa dirinya menderita diabetes [1].

Provinsi penyumbang jumlah kasus penderita diabetes melitus terbanyak di Indonesia pada tahun 2019 adalah Provinsi Jawa Timur. Jumlah penderita diabetes melitus di Jawa Timur pada tahun 2019 sebanyak 841.994 kasus [2]. Penelitian [3] mengenai faktor-faktor yang mempengaruhi diabetes melitus di Amerika Serikat menggunakan metode *Geographically Weighted Regression Forest* (GWRF) dan *Geographically Weighted Ordinary Least Square Regression* (GWOLSR). Hasil penelitian tersebut menyatakan bahwa dengan metode GWR dan GWOLS yang berpengaruh signifikan terhadap diabetes melitus adalah tingkat pendidikan, kemiskinan, obesitas, keaktifan kegiatan fisik, akses untuk berolahraga, dan indeks konsumsi bahan pangan. Kemudian penelitian [4] tentang analisis spasial mengenai asuransi kesehatan di German Timur menggunakan *Kernel Density Estimation* dan *Geographically Weighted Regression*. Hasil penelitian tersebut menghasilkan faktor-faktor yang berpengaruh signifikan adalah faktor usia, proporsi karyawan yang memiliki asuransi sosial, rata-rata pajak pendapatan, dan proporsi pasangan belum menikah.

Jumlah kasus diabetes melitus di Jawa Timur merupakan data *count* atau yang mengikuti distribusi Poisson. Regresi Poisson merupakan metode yang tepat untuk menganalisis data diskrit jika *mean* dan variansnya sama (*equidispersion*). Akan tetapi pada data jumlah kasus diabetes melitus kondisi *equidispersion* tidak terpenuhi, karena nilai varians lebih besar dari nilai *mean* (*overdispersion*). Untuk mengatasi kasus overdispersi dalam penelitian ini dilakukan pemodelan menggunakan metode *Generalized Poisson Regression* (GPR) dan *Negative Binomial Regression* (NBR). Pemodelan GPR dan

NBR menghasilkan parameter yang bersifat global untuk seluruh lokasi. Hasil interpretasi dari model ini menganggap bahwa setiap lokasi memiliki karakteristik yang sama, padahal dalam kasus diabetes melitus pada setiap lokasi memiliki karakteristik yang berbeda-beda, sehingga untuk mengatasi hal tersebut dapat digunakan analisis data spasial.

Berdasarkan uraian sebelumnya, maka pada penelitian ini dilakukan pemodelan faktor yang mempengaruhi jumlah kasus diabetes melitus di Jawa Timur menggunakan metode *Geographically Weighted Generalized Poisson Regression* (GWGPR) dan *Geographically Weighted Negative Binomial Regression* (GWNBR). GWGPR merupakan metode pengembangan dari GPR yang memperhatikan pembobot berupa letak lintang dan bujur dari titik-titik pengamatan yang diamati [5]. Metode lain yang digunakan dalam penelitian ini adalah GWNBR. Kedua metode tersebut sama-sama digunakan untuk mengatasi kasus overdispersi pada regresi Poisson dan ada penambahan efek spasial. Perbedaannya adalah terletak pada fungsi distribusinya. Pada penelitian ini dilakukan perbandingan dari kedua metode tersebut.

II. TINJAUAN PUSTAKA

A. Statistika Deskriptif

Statistika deskriptif merupakan alat yang digunakan untuk menggambarkan atau mendeskripsikan kumpulan sebuah data atau hasil pengamatan menjadi sebuah informasi [6]. Statistika deskriptif yang digunakan antara lain *mean*, varians, median, kuartil, maksimum, dan minimum.

B. Peta Tematik

Peta tematik adalah peta yang menyajikan tema tertentu dan untuk kepentingan tertentu dengan menggunakan peta rupa bumi yang telah disederhanakan sebagai dasar untuk meletakkan informasi tematiknya [7]. Pewarnaan pada peta tematik ditunjukkan untuk membedakan wilayah satu dengan wilayah lainnya. Salah satu metode yang digunakan untuk pewarnaan peta yaitu dengan metode *Natural Breaks*.

C. Multikolinearitas

Multikolinearitas adalah suatu kondisi dimana terdapat hubungan atau korelasi secara signifikan (nilainya tinggi) antara variabel prediktor dengan variabel prediktor lain yang mengakibatkan hasil taksiran parameter menjadi tidak akurat (kecil signifikansinya bahkan tidak sama sekali). Untuk mendeteksi ada atau tidaknya multikolinearitas dapat dilihat melalui nilai *Variance Inflation Factor* (VIF). Nilai VIF yang lebih besar dari 10 menunjukkan adanya multikolinearitas dimana salah satu cara mengatasinya dengan mengeluarkan variabel prediktor dengan dimulai dari yang memiliki nilai VIF terbesar [8]. Nilai VIF dapat dirumuskan oleh persamaan (1).

$$VIF_j = \frac{1}{1 - R_j^2} \tag{1}$$

dimana

$$R_j^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (\hat{x}_{ij} - \bar{x}_j)^2}{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}; j = 1, 2, \dots, k$$

Dengan R_j^2 adalah koefisien determinasi antara meregresikan satu variabel prediktor X_j dengan variabel prediktor lainnya.

D. Uji Distribusi Data

Pengujian distribusi data dilakukan untuk mengetahui distribusi probabilitas yang paling sesuai pada sebuah data. Pengujian distribusi data ada banyak namun yang paling umum digunakan adalah metode *Kolmogorov-Smirnov*. Apabila distribusi $F(y)$ akan diduga dengan $F_0(y)$, maka akan ditetapkan hipotesis nol dan pengujian yang sesuai yaitu sebagai berikut.

H_0 : $F(y) = F_0(y)$, atau data mengikuti distribusi yang ditentukan

H_1 : $F(y) \neq F_0(y)$, atau data tidak mengikuti distribusi yang ditentukan

Statistik uji *Kolmogorov-Smirnov* (D) seperti pada persamaan (2) [9].

$$D = \max |(S(y) - F_0(y))| \tag{2}$$

keterangan:

$S(y)$ = nilai empiris distribusi kumulatif sampel

$F_0(y)$ = fungsi distribusi kumulatif yang ditentukan

D = nilai kritis uji *Kolmogorov-Smirnov*

Daerah penolakan H_0 apabila $D_{hitung} > D_{(1-\alpha, n)}$ atau $p\text{-value} < \alpha$, yang artinya distribusi probabilitas yang diduga tidak sama dengan distribusi tertentu.

E. Regresi Poisson

Regresi Poisson merupakan model regresi non-linier yang tepat digunakan untuk memodelkan data cacahan. Fungsi probabilitas variabel random diskrit (y) berdistribusi Poisson dengan parameter μ dinyatakan sebagaimana persamaan (3).

$$f(y, \mu) = \frac{e^{-\mu} \mu^y}{y!}; y = 0, 1, 2, \dots; \mu > 0 \tag{3}$$

dimana μ merupakan rata-rata banyaknya kejadian dalam suatu interval tertentu. Nilai harapan dan ragam dari distribusi Poisson adalah $E(Y) = \text{Var}(Y) = \mu$. Model regresi Poisson dapat ditulis sebagai berikut.

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \tag{4}$$

dengan \mathbf{x} adalah vektor variabel prediktor dan $\boldsymbol{\beta}$ adalah parameter regresi Poisson yang dinotasikan sebagai berikut:

$$\begin{aligned} \mathbf{x}_i &= [1 \ x_{i1} \ x_{i2} \ \dots \ x_{ik}]^T \\ \boldsymbol{\beta} &= [\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_k]^T \end{aligned} \tag{5}$$

F. Overdispersi

Regresi Poisson memiliki asumsi yaitu ekuidispersi dimana variabel respon yang digunakan memiliki nilai *mean* dan varians yang sama. Salah satu pelanggaran asumsi tersebut adalah overdispersi dimana nilai varians lebih besar dari nilai *mean*. Overdispersi menyebabkan penaksir dari parameter regresinya tetap konsisten namun tidak efisien. Hal ini berdampak pada nilai *standar error* yang menjadi *underestimate*, sehingga kesimpulan yang diperoleh menjadi tidak valid [10]. Overdispersi merupakan nilai *deviance* yang dibagi derajat bebasnya diperoleh nilai lebih besar daripada 1. Pemeriksaan overdispersi dapat dilakukan sebagaimana persamaan (6).

$$\hat{\theta} = \frac{\text{Deviance}}{db}; db = n - k \tag{6}$$

dimana

$$\text{Deviance} = 2 \sum_{i=1}^n \left(y_i \ln \left(\frac{y_i}{\hat{y}_i} \right) - (y_i - \hat{y}_i) \right)$$

Jika nilai dari $\hat{\theta} > 1$ artinya terdapat overdispersi pada model regresi Poisson, jika $\hat{\theta} < 1$ artinya terdapat underdispersi dan jika $\hat{\theta} = 1$ berarti tidak terjadi kasus over/under dispersi yang disebut ekuidispersi.

G. Uji Heterogenitas Spasial

Heterogenitas spasial adalah variasi pada tiap lokasi dengan tipe data spasial yaitu spasial titik. Uji heterogenitas spasial dilakukan karena setiap daerah mempunyai ciri khas yang berbeda dengan daerah lainnya. Uji ini dilakukan dengan *Breusch-Pagan* (BP) dengan hipotesisnya adalah sebagai berikut.

- $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$ (varians antar lokasi sama)
- $H_1: \text{minimal ada satu } \sigma_i^2 \neq \sigma^2; i=1, 2, \dots, n$ (varians antar lokasi berbeda)

Rumus statistik uji sebagaimana ditunjukkan pada persamaan (7).

$$BP = \left(\frac{1}{2} \right) \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f} \sim \chi_{(k)}^2 \tag{7}$$

dimana,

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

$e_i = (y_i - \hat{y}_i)$, merupakan residual pengamatan ke- i

$\hat{\sigma}^2 = \text{varians sampel } y$

$\mathbf{Z} = \text{matriks ukuran } n \times (k + 1)$ berisi vektor yang dinormalkan (z) untuk setiap pengamatan, dengan k adalah banyaknya prediktor.

Daerah penolakan H_0 jika nilai $BP > \chi_{(\alpha; k)}^2$ atau jika $p\text{-value} < \alpha$, yang berarti variansi antar lokasi berbeda.

H. Matriks Pembobot Spasial

Keragaman spasial antara lokasi satu dengan lokasi lain ditunjukkan dengan adanya matriks pembobot \mathbf{W} . Elemen dari matriks pembobot spasial (\mathbf{W}) merupakan fungsi jarak *eulidian* antar lokasi. Salah satu fungsi pembobot adalah fungsi *adaptive bisquare* kernel. Fungsi *adaptive bisquare* kernel merupakan fungsi kernel dengan *bandwidth* yang berbeda pada setiap lokasi pengamatan [11]. Fungsi *Adaptive Bisquare* Kernel dapat dirumuskan sebagaimana persamaan (8).

$$w_{ii^*} = \begin{cases} \left(1 - \left(\frac{d_{ii^*}}{h_i} \right)^2 \right)^2 & ; \text{untuk } d_{ii^*} \leq h_i \\ 0 & ; \text{untuk } d_{ii^*} \geq h_i \end{cases} \tag{8}$$

dimana

$$d_{ii^*} = \sqrt{(u_i - u_{i^*})^2 + (v_i - v_{i^*})^2}$$

d_{ii^*} adalah jarak *Euclidian* antara lokasi ke- i dan lokasi ke- i^* . Sedangkan h_i adalah parameter penghalus atau yang disebut sebagai *bandwidth* dari lokasi ke- i . Pemilihan *bandwidth* optimum diperlukan untuk mengatur varian dan bias dari penaksir yang dihasilkan. Pemilihan *bandwidth* optimum dapat dilakukan menggunakan metode *Cross Validation* (CV). Metode CV ini didefinisikan oleh persamaan (9) [12].

$$CV(h_i) = \sum_{i=1}^n (Y_i - \hat{Y}_{zi}(h_i))^2 \tag{9}$$

I. Generalized Poisson Regression

Model GPR merupakan model yang sesuai untuk data *count* apabila terjadi *over/under* dispersi. Sehingga selain parameter μ , dalam model GPR juga terdapat θ sebagai parameter dispersi. Misal $y = 0,1,2,\dots$ merupakan variabel respon. Distribusi GPR [13] adalah sebagaimana persamaan (10).

$$f(y, \mu, \theta) = \left(\frac{\mu}{1 + \theta\mu}\right)^y \frac{(1 + \theta y)^{y-1}}{y!} \exp\left(\frac{-\mu(1 + \theta y)}{1 + \theta\mu}\right), \tag{10}$$

$y = 0,1,2,\dots; \mu > 0; -\infty < \theta < \infty$

Mean dan varians model GPR adalah $E(y) = \mu$ dan $var(y) = \mu(1 + \theta\mu)^2$. Apabila θ sama dengan 0 maka model GPR akan menjadi model regresi Poisson biasa, sedangkan apabila $\theta > 0$ maka model GPR merepresentasikan terjadi *overdispersion* pada data *count*, dan apabila $\theta < 0$ maka model GPR merepresentasikan terjadi *underdispersion* pada data *count*. Model GPR memiliki bentuk sama dengan model regresi Poisson seperti pada persamaan (11).

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \tag{11}$$

J. Geographically Weighted Generalized Poisson Regression (GWGPR)

GWGPR merupakan pengembangan dari GPR. Model GWGPR menggunakan pembobot geografis pada penaksiran parameternya. Sehingga model GWGPR akan menghasilkan penaksir parameter lokal, dengan parameter yang berbeda untuk setiap lokasi [14]. Fungsi distribusi GWGPR untuk setiap lokasi adalah sebagaimana persamaan (12) [5].

$$f(y_i, \mu_i, \theta_i) = \left(\frac{\mu_i}{1 + \theta_i\mu_i}\right)^{y_i} \frac{(1 + \theta_i y_i)^{y_i-1}}{y_i!} \exp\left(\frac{-\mu_i(1 + \theta_i y_i)}{1 + \theta_i\mu_i}\right) \tag{12}$$

Bentuk persamaan GWGPR adalah seperti pada persamaan (2.13).

$$\mu_i = e^{\mathbf{x}_i^T \boldsymbol{\beta}(u_i, v_i)} = \exp(\beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \dots + \beta_k(u_i, v_i)x_{ik}) \tag{13}$$

Metode yang digunakan untuk menaksir parameter model GWGPR adalah *Maximum Likelihood Estimation* (MLE), yaitu dengan memaksimalkan faktor fungsi *likelihood* [15]. Fungsi logaritma natural (*ln-likelihood*) berdasarkan faktor geografis seperti pada persamaan (14).

$$\begin{aligned} \ln(L^*(\beta(u_{i^*}, v_{i^*}, \theta_i))) &= \sum_{i=1}^n w_{i^*} \left(y_i \left(\ln e^{\mathbf{x}_i^T \boldsymbol{\beta}(u_{i^*}, v_{i^*})} - \ln(1 + \theta_i e^{\mathbf{x}_i^T \boldsymbol{\beta}(u_{i^*}, v_{i^*})}) \right) \right. \\ &\quad \left. + (y_i - 1) \ln(1 + \theta_i y_i) - \ln y_i! - \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}(u_{i^*}, v_{i^*})} (1 + \theta_i y_i)}{1 + \theta_i e^{\mathbf{x}_i^T \boldsymbol{\beta}(u_{i^*}, v_{i^*})}} \right) \\ &= \sum_{i=1}^n w_{i^*} \left(y_i \left(\mathbf{x}_i^T \boldsymbol{\beta}(u_{i^*}, v_{i^*}) \right) - \ln(1 + \theta_i e^{\mathbf{x}_i^T \boldsymbol{\beta}(u_{i^*}, v_{i^*})}) \right) \\ &\quad \left. + (y_i - 1) \ln(1 + \theta_i y_i) - \ln(y_i!) - \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}(u_{i^*}, v_{i^*})} (1 + \theta_i y_i)}{1 + \theta_i e^{\mathbf{x}_i^T \boldsymbol{\beta}(u_{i^*}, v_{i^*})}} \right) \end{aligned} \tag{14}$$

Proses mendapatkan penaksir parameter dari model GWGPR adalah dengan menurunkan persamaan (14) terhadap masing-masing parameternya kemudian disamakan dengan nol. Namun hasilnya tidak dapat dilakukan secara analitik karena hasil yang didapat tidak *close form*, sehingga perlu dilakukan iterasi yaitu dengan iterasi *Newton Raphson*.

Pengujian signifikansi model GWGPR terdiri dari uji serentak dan uji parsial. Salah satu pengujian parameter model secara serentak dapat menggunakan *Maximum Likelihood Ratio Test* (MLRT) untuk menguji signifikansi dari faktor geografis atau lokasi. Hipotesis yang digunakan adalah sebagai berikut [16].

$$H_0 : \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_k(u_i, v_i) = 0 ; i = 1, 2, \dots, n$$

$$H_1 : \text{minimal ada satu } \beta_j(u_i, v_i) \neq 0 ; i = 1, 2, \dots, n ; j = 1, 2, \dots, k$$

Statistik uji yang digunakan adalah sebagaimana persamaan (15).

$$G^2 = -2 \ln \hat{\Lambda} = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = 2 \left(\ln L(\hat{\Omega}) - \ln L(\hat{\omega}) \right) \tag{15}$$

$L(\hat{\omega})$ merupakan fungsi *likelihood* di bawah H_0 untuk model sederhana tanpa melibatkan variabel x dan $L(\hat{\Omega})$ merupakan fungsi *likelihood* di bawah H_1 untuk model dengan melibatkan variabel x . Daerah penolakan H_0 jika nilai dari $G^2 > \chi^2_{(\alpha, k)}$ yang artinya minimal terdapat satu parameter dalam model GWGPR yang berpengaruh signifikan terhadap variabel respon.

Setelah melakukan pengujian secara serentak, maka dilanjutkan dengan pengujian parameter model secara parsial untuk mengetahui parameter mana saja yang signifikan mempengaruhi variabel responnya. Berikut merupakan hipotesis yang digunakan.

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

$$H_1 : \beta_j(u_i, v_i) \neq 0; j = 1, 2, \dots, k$$

Statistik uji yang digunakan adalah seperti pada persamaan (16).

$$Z = \frac{\hat{\beta}_j(u_i, v_i)}{se(\hat{\beta}_j(u_i, v_i))} \tag{16}$$

dimana $se(\hat{\beta}_j(u_i, v_i)) = \sqrt{\widehat{var}(\hat{\beta}_j(u_i, v_i))}$, dengan $\widehat{var}(\hat{\beta}_j(u_i, v_i))$ merupakan ekspektasi diagonal ke (j+1) pada matriks kovarians dari $\hat{\beta}$ yang merupakan negatif invers matriks Hessien $H^{-1}(\hat{\beta}_{(m)})$. Daerah penolakan H_0 adalah jika nilai statistik uji $|Z_{hitung}| > Z_{(\alpha/2)}$ yang artinya variabel prediktor tersebut berpengaruh signifikan terhadap variabel respon pada setiap lokasi dalam model GWGPR [15].

K. Negative Binomial Regression

Regresi binomial negatif merupakan salah satu solusi untuk mengatasi adanya kasus overdispersi pada model regresi Poisson. Model regresi binomial negatif memiliki fungsi massa peluang seperti pada persamaan (17) [17].

$$f(y, \mu, \theta) = \frac{\Gamma(y + \theta^{-1})}{\Gamma(\theta^{-1})\Gamma(y + 1)} \left(\frac{1}{1 + \theta\mu}\right)^{\theta^{-1}} \left(\frac{\theta\mu}{1 + \theta\mu}\right)^y, \tag{17}$$

$y = 0, 1, 2, \dots; \mu > 0; \theta > 0$

Mean dan varians model NBR adalah $E(y) = \mu$ dan $var(y) = \mu + \theta\mu^2$. Model binomial negatif merupakan paduan antara distribusi Poisson dan Gamma [18]. Berikut merupakan model regresi binomial negatif.

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik}) \tag{18}$$

Fungsi peluang dari distribusi binomial negatif adalah sebagaimana persamaan (2.19).

$$\ln f(y, \mu, \theta) = \ln \frac{\Gamma(y + \theta^{-1})}{\Gamma(\theta^{-1})\Gamma(y + 1)} + \frac{1}{\theta} \ln \left(\frac{1}{1 + \theta\mu}\right) + y \ln \left(\frac{\theta\mu}{1 + \theta\mu}\right) \tag{19}$$

Pendugaan parameter model regresi binomial negatif menggunakan metode MLE. Estimasi regresi binomial negatif menggunakan metode iterasi *Newton Raphson* untuk memaksimumkan fungsi *likelihood*. Fungsi *likelihood* dari regresi binomial negatif adalah seperti pada persamaan (20).

$$L(\beta, \theta) = \prod_{i=1}^n \left\{ \frac{\Gamma(y_i + \theta^{-1})}{\Gamma(\theta^{-1})\Gamma(y_i + 1)} \left(\frac{1}{1 + \theta e^{x_i^T \beta}}\right)^{\theta^{-1}} \left(\frac{\theta e^{x_i^T \beta}}{1 + \theta e^{x_i^T \beta}}\right)^{y_i} \right\} \tag{20}$$

Bentuk *ln-likelihood* adalah seperti pada persamaan (2.21).

$$\begin{aligned} \ln L(\beta, \theta) &= \sum_{i=1}^n \left\{ \ln \frac{\Gamma(y_i + \theta^{-1})}{\Gamma(\theta^{-1})\Gamma(y_i + 1)} + \theta^{-1} \ln \left(\frac{1}{1 + \theta e^{x_i^T \beta}}\right) + y_i \ln \left(\frac{\theta e^{x_i^T \beta}}{1 + \theta e^{x_i^T \beta}}\right) \right\} \\ &= \sum_{i=1}^n \left\{ \ln \frac{\Gamma(y_i + \theta^{-1})}{\Gamma(\theta^{-1})\Gamma(y_i + 1)} + y_i \ln(\theta e^{x_i^T \beta}) - (y_i + \theta^{-1}) \ln(1 + \theta e^{x_i^T \beta}) \right\} \end{aligned} \tag{21}$$

L. Geographically Weighted Negative Binomial Regression (GWNBR)

Pendugaan GWNBR merupakan salah satu metode yang cukup efektif menduga data yang memiliki heterogenitas spasial untuk data *count* yang overdispersi. Model GWNBR merupakan pengembangan dari model regresi binomial negatif. Model GWNBR menghasilkan parameter lokal dengan masing-masing lokasi akan memiliki parameter yang berbeda-beda. Model GWNBR dapat dirumuskan sebagaimana persamaan (22) [19].

$$y_i \sim NB \left[\exp \left(\sum_{j=0}^k \beta_j(u_i, v_i) x_{ij} \right), \theta_i \right], i = 1, 2, \dots, n; x_{i0} = 1 \tag{22}$$

dimana,

- y_i : nilai observasi respon ke- i
- x_{ij} : nilai observasi variabel prediktor ke- j pada pengamatan lokasi (u_i, v_i)
- $\beta_j(u_i, v_i)$: koefisien regresi variabel prediktor ke- j untuk setiap lokasi (u_i, v_i)
- θ_i : parameter dispersi untuk setiap lokasi (u_i, v_i)

Fungsi distribusi binomial negatif untuk setiap lokasi dapat dituliskan sebagaimana persamaan berikut.

$$f(y_i, \beta(u_i, v_i), \theta_i) = \frac{\Gamma\left(y_i + \frac{1}{\theta_i}\right)}{\Gamma\left(\frac{1}{\theta_i}\right)\Gamma(y_i + 1)} \left(\frac{1}{1 + \theta_i e^{x_i^T \beta(u_i, v_i)}}\right)^{\frac{1}{\theta_i}} \left(\frac{\theta_i e^{x_i^T \beta(u_i, v_i)}}{1 + \theta_i e^{x_i^T \beta(u_i, v_i)}}\right)^{y_i}, i = 1, 2, \dots, n \tag{23}$$

Bentuk persamaan GWNBR adalah sebagai berikut.

$$\mu_i = e^{x_i^T \beta(u_i, v_i)} = \exp(\beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \dots + \beta_k(u_i, v_i)x_{ik}) \tag{24}$$

Estimasi model GWNBR menggunakan MLE. Fungsi Likelihood sebaran binomial negatif untuk tiap lokasi adalah sebagai berikut.

$$L(\beta(u_i, v_i), \theta, i = 1, 2, \dots, n) = \prod_{i=1}^n \left(\frac{\Gamma(y_i + \frac{1}{\theta_i})}{\Gamma(\frac{1}{\theta_i}) \Gamma(y_i + 1)} \left(\frac{1}{1 + \theta_i e^{\beta(u_i, v_i)}} \right)^{\frac{1}{\theta_i}} \left(\frac{\theta_i e^{\beta(u_i, v_i)}}{1 + \theta_i e^{\beta(u_i, v_i)}} \right)^{y_i} \right) \quad (25)$$

Proses mendapatkan penaksir parameter dari model GWNBR hasilnya tidak dapat dilakukan secara analitik karena hasil yang didapat tidak *close form*, sehingga perlu dilakukan iterasi yaitu dengan iterasi *Newton Raphson*.

Pengujian signifikansi model GWNBR terdiri dari uji serentak dan uji parsial. Salah satu pengujian parameter model secara serentak dapat menggunakan *Maximum Likelihood Ratio Test* (MLRT). Pengujian secara serentak dan parsial dilakukan sama persis seperti pada pengujian di GWGPR.

M. Kriteria Keباikan Model

Model terbaik adalah model yang mampu menjelaskan hubungan antara variabel prediktor dengan variabel respon berdasarkan kriteria tertentu. Pemilihan model terbaik menggunakan *Akaike's Information Criteria Corrected* (AICc). AICc merupakan pengembangan dari AIC dengan mempertimbangkan banyak parameter dan banyak data. Apabila semakin banyak parameter yang digunakan, maka AICc lebih baik daripada AIC. Rumus dari AICc seperti pada persamaan (26) [20].

$$AICc = -2 \ln \left(L(\beta(u_i, v_i), i = 1, 2, \dots, n) \right) + 2k + \frac{2k(k+1)}{n-k-1} \quad (26)$$

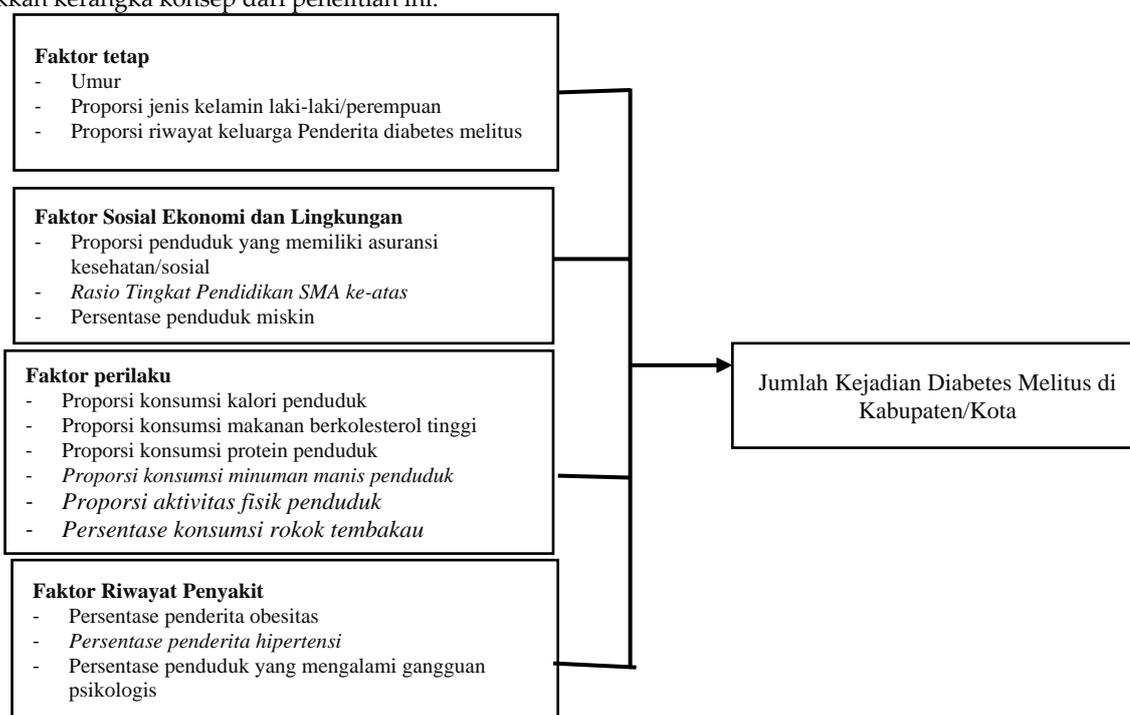
dengan *n* adalah banyak pengamatan dan *k* adalah jumlah parameter yang diestimasi dalam model.

N. Diabetes Melitus

Diabetes melitus merupakan suatu penyakit metabolik yang ditandai dengan hiperglikemia (kondisi kadar gula darah yang tinggi) yang terjadi karena kelainan sekresi insulin, kerja insulin atau keduanya [21]. Diabetes adalah masalah kesehatan masyarakat yang penting, menjadi salah satu dari empat penyakit tidak menular prioritas yang menjadi target tindak lanjut oleh para pemimpin dunia. Jumlah kasus dan prevalensi diabetes terus meningkat selama beberapa dekade terakhir [22]. Diabetes melitus dibedakan menjadi tiga jenis, yang pertama yaitu diabetes melitus tipe 1. Diabetes melitus tipe 1 merupakan diabetes yang disebabkan kenaikan kadar gula darah yang disebabkan kerusakan sel beta pankreas sehingga produksi insulin tidak ada sama sekali. Kedua, diabetes melitus tipe 2 adalah diabetes yang disebabkan oleh kenaikan gula darah karena penurunan sekresi insulin yang rendah oleh kelenjar pankreas. Ketiga adalah diabetes melitus tipe gestasional. Diabetes tipe ini ditandai dengan kenaikan kadar gula darah selama masa kehamilan. Gangguan ini biasa terjadi pada minggu ke-24 kehamilan dan kadar gula darah akan kembali normal setelah persalinan [1].

O. Kerangka Konseptual

Kerangka konseptual pada penelitian ini merujuk pada faktor penyebab kasus diabetes melitus. Gambar 1 menunjukkan kerangka konsep dari penelitian ini.



Keterangan:
Variabel yang diteliti adalah yang bercetak miring.

Gambar 1 Kerangka Konsep Penelitian

III. METODOLOGI

A. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder tahun 2019 yang penulis dapatkan melalui Dinas Kesehatan Provinsi Jawa Timur tahun dan Badan Pusat Statistik. Data yang digunakan mencakup data mengenai penderita diabetes melitus di Provinsi Jawa Timur dan faktor-faktor yang diduga mempengaruhinya. Unit pengamatan yang digunakan sebanyak 38 Kabupaten/Kota di Jawa Timur.

B. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini dijelaskan pada Tabel 1.

Tabel 1 Variabel Penelitian

Indeks	Variabel	Skala Data
Y	Jumlah kasus Diabetes Melitus	Rasio
X ₁	Rasio tingkat pendidikan SMA ke-atas	Rasio
X ₂	Proporsi aktivitas fisik penduduk	Rasio
X ₃	Persentase konsumsi rokok tembakau	Rasio
X ₄	Persentase penderita hipertensi	Rasio
X ₅	Proporsi konsumsi minuman manis	Rasio

C. Langkah Analisis

Berikut merupakan langkah-langkah untuk analisis faktor-faktor yang mempengaruhi jumlah kasus diabetes melitus di Provinsi Jawa Timur.

1. Mendeskripsikan jumlah kasus diabetes melitus serta variabel-variabel yang diduga mempengaruhi menggunakan statistika deskriptif yaitu *mean*, standar deviasi, maksimum, dan minimum.
2. Membuat peta tematik pada kabupaten/kota di Provinsi Jawa Timur menggunakan metode *Neural Breaks*.
3. Menguji distribusi dari variabel respon.
4. Menguji multikolinearitas antar variabel prediktor berdasarkan persamaan (2.1).
5. Menganalisis model regresi Poisson dan overdispersi.
6. Melakukan pemodelan dengan GPR dan Regresi Binomial Negatif untuk jumlah kasus diabetes melitus.
7. Melakukan pengujian aspek data spasial dengan melakukan uji *Breusch-Pagan* untuk melihat heterogenitas spasial berdasarkan persamaan (2.7).
8. Mendapatkan model GWGPR dan GWNBR untuk jumlah kasus diabetes melitus beserta faktor-faktor yang diduga berpengaruh.
9. Membandingkan nilai AICc diantara model GWGPR dan GWNBR.
10. Melakukan interpretasi pada model yang terbaik.
11. Menarik kesimpulan dan saran.

IV. ANALISIS DAN PEMBAHASAN

A. Karakteristik Data

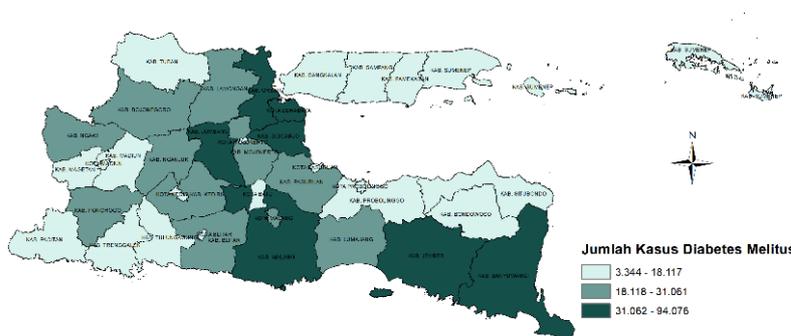
Sebagaimana telah dijelaskan pada bab 3 bahwa penelitian ini menggunakan data jumlah kasus diabetes melitus yang terjadi di 38 Kabupaten/Kota di Jawa Timur tahun 2019 dengan faktor-faktor yang diduga mempengaruhi. Jumlah penderita diabetes melitus tertinggi di Indonesia berada di Jawa Timur tahun 2019 yaitu sebanyak 841.994 jiwa. Berikut merupakan gambaran awal tentang jumlah kasus diabetes melitus di Jawa Timur dengan peta tematik.

Hasil analisis statistika deskriptif pada data jumlah kasus diabetes melitus di Jawa Timur tahun 2019. Statistika deskriptif yang digunakan seperti *mean*, standar deviasi, minimum, dan maksimum seperti pada Tabel 2.

Tabel 2 Statistika Deskriptif Variabel Penelitian

Variabel	Mean	StDev	Minimum	Maksimum
Y	22.158	18.273	3.344	94.076

Berdasarkan Tabel 2 diperoleh informasi bahwa rata-rata jumlah penderita diabetes melitus di Jawa Timur tahun 2019 adalah sebesar 22.158, dengan standar deviasi sebesar 18.273. Hal tersebut menandakan bahwa persebaran yang sangat besar pada jumlah kasus diabetes melitus di kabupaten/kota Jawa Timur dimana terdapat data yang ribuan hingga puluhan ribu. Daerah yang memiliki jumlah kasus diabetes melitus terkecil ada di Kota Batu dengan jumlah kasus sebesar 3.344 orang, sedangkan daerah dengan jumlah penderita diabetes melitus terbanyak ada di Kota Surabaya dengan jumlah penderita sebanyak 94.076 orang.



Gambar 2 Persebaran Jumlah Kasus Diabetes Melitus Menurut Kabupaten/Kota di Provinsi Jawa Timur

Jumlah kasus diabetes melitus di Jawa Timur terus meningkat hingga tahun 2019. Jumlah penderita diabetes melitus tahun 2019 adalah 841.994 orang. Peta persebaran Gambar 2 menunjukkan pola yang mengelompok untuk jumlah penderita diabetes di Kabupaten/Kota. Pada peta di atas terdapat kelompok kabupaten/kota yang diabetes melitus tergolong tinggi dengan rentang 31.062 hingga 94.076 yaitu 8 Kabupaten/Kota di Jawa Timur yaitu Kabupaten Malang, Jember, Banyuwangi, Sidoarjo, Mojokerto, Jombang, Gresik, dan Surabaya. Daerah yang memiliki jumlah kasus diabetes melitus dengan kategori sedang antara lain Kabupaten Ponorogo, Blitar, Kediri, Lumajang, Probolinggo, Pasuruan, Nganjuk, Ngawi, Bojonegoro, Lamongan, dan Kota Malang. Sedangkan daerah dengan kategori rendah pada jumlah kasus diabetes melitus menduduki peringkat terbanyak yaitu di 19 kabupaten/kota di Jawa Timur dengan rentang nilai 31.062 hingga 94.076 kasus.

B. Korelasi Antar Variabel

Jumlah kas Korelasi digunakan untuk mengetahui arah hubungan dan kekuatan antar dua variabel yang digunakan dengan melihat besarnya nilai korelasi *Pearson* yang dihasilkan. Nilai korelasi antar variabel ditampilkan pada Tabel 3.

Tabel 3 Korelasi Antar Variabel

Variabel	Y	X ₁	X ₂	X ₃	X ₄
X ₁	0,324 0,047*				
X ₂	-0,136 0,415*	-0,329 0,044*			
X ₃	-0,132 0,429*	-0,608 0,000*	0,239 0,148*		
X ₄	-0,001 0,996*	-0,098 0,558*	0,229 0,167*	0,139 0,407*	
X ₅	0,105 0,529*	-0,042 0,800*	-0,362 0,025*	0,045 0,789*	-0,14 0,401*

Berdasarkan Tabel 3 dapat diketahui bahwa hubungan antara jumlah kasus diabetes melitus di Provinsi Jawa Timur tahun 2019 dengan lima variabel prediktor lainnya secara umum memiliki nilai korelasi yang cukup rendah karena memiliki nilai korelasi di bawah 0,7. Hubungan antara jumlah kasus diabetes melitus (Y) dengan proporsi aktivitas fisik (X₂), persentase konsumsi rokok (X₃), dan persentase penderita hipertensi (X₄) memiliki hubungan yang negatif. Sedangkan untuk variabel rasio tingkat pendidikan SMA ke-atas (X₁) dan proporsi konsumsi minuman manis (X₅) memiliki hubungan yang positif dengan variabel Y.

C. Pemeriksaan Multikolinearitas

Berdasarkan Pengujian multikolinearitas dilakukan untuk memenuhi asumsi pada pemodelan regresi, dimana antar variabel harus independen atau tidak ada hubungan antar variabel prediktor. Adanya multikolinearitas dapat diketahui dari nilai VIF setiap variabel prediktor sebagaimana seperti pada persamaan (2.1). Nilai VIF yang kecil menunjukkan tidak adanya kasus multikolinearitas antar variabel prediktor. Nilai VIF pada setiap variabel prediktor ditampilkan pada Tabel 4.

Tabel 4 Nilai VIF dari Variabel Prediktor

Variabel	VIF
X ₁	1,7087
X ₂	1,3750
X ₃	1,6091
X ₄	1,0707
X ₅	1,2007

Tabel 4 menginformasikan bahwa semua variabel prediktor telah memenuhi asumsi non multikolinearitas karena nilai VIF dari setiap variabel predictor sangat kecil. Hal ini menunjukkan bahwa tidak ada variabel prediktor yang saling berkorelasi dengan variabel prediktor lainnya. Sehingga lima variabel prediktor dapat digunakan untuk pembentukan model regresi poisson, NBR, GPR, GWGPR, dan GWNBR.

D. Pengujian Distribusi Data

Pada penelitian ini dilakukan perbandingan dari hasil model GWGPR dan GWNBR dimana data variabel respon harus berdistribusi *Generalized Poisson* atau *Negative Binomial*. Berdasarkan hasil analisis didapatkan nilai *p-value* sebesar 0,00336 untuk *Negative Binomial* dan *p-value* sebesar $3,4473 \cdot 10^{-14}$ untuk distribusi Poisson dimana nilai tersebut lebih kecil dari signifikansi 5% sehingga keputusan yang diambil adalah tolak H_0 dan kesimpulan yang diambil adalah data jumlah kasus diabetes melitus tidak berdistribusi keduanya.

E. Pemeriksaan Overdispersi

Pemeriksaan overdispersi dilakukan berdasarkan hasil dari uji serentak dari regresi Poisson antar variabel respon dengan variabel prediktor. Tabel 5 menunjukkan hasil estimasi parameter model regresi Poisson.

Tabel 5 Nilai Estimasi Regresi Poisson

Parameter	Estimasi	SE($\hat{\beta}$)	Z hitung	P-value
$\hat{\beta}_0$	8,2970000	0,0201100	412,49	< 2.10 ⁻¹⁶ *
$\hat{\beta}_1$	0,0270700	0,0001248	216,85	< 2.10 ⁻¹⁶ *
$\hat{\beta}_2$	0,0001058	0,0001722	0,6150	0,539
$\hat{\beta}_3$	0,0273500	0,0005418	50,470	< 2.10 ⁻¹⁶ *
$\hat{\beta}_4$	0,0032160	0,0000939	34,241	< 2.10 ⁻¹⁶ *
$\hat{\beta}_5$	0,0324800	0,0003968	81,859	< 2.10 ⁻¹⁶ *
Devians	382.181			

*) signifikan dengan taraf 5%

Regresi Poisson memiliki ciri bahwa nilai *mean* dan varians yang disebut ekuidispersi, tetapi kasus jumlah diabetes melitus terjadi overdispersi. Untuk mendeteksi keberadaan overdispersi adalah nilai devians model regresi Poisson sebesar 382.181 dengan derajat bebas 32, sehingga rasio nilai devians dengan derajat bebasnya bernilai 11.943,156. Nilai tersebut lebih besar daripada 1 yang artinya data jumlah kasus diabetes melitus mengalami overdispersi. Regresi poisson tidak tepat untuk kasus overdispersi karena akan menghasilkan estimasi parameter yang tidak efisien dan bias

F. Pemodelan GPR Jumlah Kasus Diabetes Melitus

Regresi Poisson Model GPR digunakan apabila terjadi kasus *over/under* dispersi pada model regresi poisson seperti pada subab (I). Hasil estimasi model GPR disajikan dalam Tabel 6.

Tabel 6 Nilai Estimasi Parameter Model GPR

Parameter	Estimasi	SE($\hat{\beta}$)	Z hitung	p-value
$\hat{\beta}_0$	4,8658435	2,4023606	2,025	0,0428*
$\hat{\beta}_1$	0,0399065	0,0194404	2,053	0,0401*
$\hat{\beta}_2$	0,0254091	0,0192263	1,322	0,1863
$\hat{\beta}_3$	0,0762239	0,0652660	1,168	0,2428
$\hat{\beta}_4$	0,0001695	0,0100325	0,017	0,9865
$\hat{\beta}_5$	0,0566466	0,0474391	1,194	0,2324
Devians	3025,006			

*) signifikan dengan taraf 5%

Dari hasil analisis pada Tabel 6 dilakukan pengujian secara serentak dan parsial. Pengujian secara serentak dilakukan untuk mengetahui apakah secara serentak variabel prediktor memberikan pengaruh terhadap variabel respon. Berikut adalah hipotesis dari pengujian GPR secara serentak.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_5 = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0 ; j = 1,2,3,4,5$$

Nilai devians yang ditunjukkan Tabel 6 adalah sebesar 3025,006. Taraf signifikansi yang digunakan yaitu sebesar 5% ($\alpha = 5\%$), sehingga nilai $\chi^2_{(0,05;5)}$ sebesar 11,0705. Hal ini menunjukkan bahwa nilai devians lebih besar dari nilai $\chi^2_{(0,05;5)}$, sehingga diperoleh keputusan Tolak H_0 yang berarti bahwa minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap model dan dilanjutkan ke uji parsial dengan hipotesis sebagai berikut.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 ; j = 1,2,3,4,5$$

Pengujian ini menggunakan statistik uji $|Z_{hitung}|$ yang kemudian dibandingkan dengan $Z_{(\alpha/2)}$. Berdasarkan hasil pengujian secara parsial dengan taraf signifikansi 5% didapatkan $Z_{(0,05/2)}$ sebesar 1,96 maka dapat disimpulkan bahwa ada satu variabel yang prediktor memberikan pengaruh signifikan secara parsial terhadap jumlah kasus diabetes melitus di Jawa Timur yaitu rasio tingkat pendidikan SMA ke-atas. Model GPR yang diperoleh adalah sebagai berikut.

$$\hat{\mu} = \exp(4,8658435 + 0,0399065X_1 + 0,0254091X_2 + 0,0762239X_3 + 0,0001695X_4 + 0,0566466X_5)$$

G. Pemodelan Regresi Binomial Negatif

Selain menggunakan metode GPR untuk mengatasi overdispersi pada model regresi Poisson, dapat juga menggunakan regresi binomial negatif. Hasil estimasi parameter model regresi binomial negatif adalah sebagai berikut.

Tabel 7 Nilai Estimasi Parameter Model Regresi Binomial Negatif

Parameter	Estimasi	SE($\hat{\beta}$)	Z hitung	P-value
$\hat{\beta}_0$	7,022833	2,039066	3,444	0,000573*
$\hat{\beta}_1$	0,030911	0,014126	2,188	0,028651*
$\hat{\beta}_2$	0,010233	0,016949	0,604	0,545996
$\hat{\beta}_3$	0,043347	0,053940	0,804	0,421614
$\hat{\beta}_4$	0,002381	0,008802	0,271	0,786742
$\hat{\beta}_5$	0,041679	0,038850	1,073	0,283353
Devians	40,757			

*) signifikan dengan taraf 5%

Berdasarkan Tabel 7 di atas dilakukan pengujian secara serentak dan parsial. Pengujian secara serentak dilakukan untuk mengetahui apakah secara serentak variabel prediktor memberikan pengaruh terhadap variabel respon. Berikut adalah hipotesis dari pengujian regresi binomial negatif secara serentak.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_5 = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0 ; j = 1,2,3,4,5$$

Nilai devians yang ditunjukkan Tabel 7 adalah sebesar 40,757. Taraf signifikansi yang digunakan yaitu sebesar 5% ($\alpha = 5\%$), sehingga nilai $\chi^2_{(0,05;5)}$ sebesar 11,0705. Hal ini menunjukkan bahwa nilai devians lebih besar dari nilai $\chi^2_{(0,05;5)}$. Sehingga diperoleh keputusan Tolak H_0 yang berarti bahwa minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap model dan dilanjutkan ke uji parsial dengan hipotesis sebagai berikut.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 ; j = 1,2,3,4,5$$

Pengujian ini menggunakan statistik uji $|Z_{hitung}|$ yang kemudian dibandingkan dengan $Z_{(\alpha/2)}$. Berdasarkan hasil pengujian secara parsial dengan taraf signifikansi 5% didapatkan $Z_{(0,05/2)}$ sebesar 1,96 maka dapat disimpulkan bahwa variabel yang berpengaruh secara signifikan adalah rasio tingkat pendidikan SMA ke-atas (X_1). Model regresi binomial negatif adalah sebagai berikut.

$$\hat{\mu} = \exp(7,022833 + 0,030911X_1 + 0,010233X_2 + 0,043347X_3 + 0,002381X_4 + 0,041679X_5)$$

H. Pengujian Aspek Data Spasial

Pengujian heterogenitas spasial dilakukan untuk mengetahui adanya keragaman spasial pada data jumlah kasus diabetes melitus di Provinsi Jawa Timur. Adanya perbedaan karakteristik antara satu titik pengamatan dengan titik pengamatan yang lain dapat dilihat dengan pengujian Uji *Breusch-Pagan* dengan hipotesis sebagai berikut.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{38}^2 = \sigma^2$$

$$H_1 : \text{minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1,2, \dots, 38$$

Didapatkan nilai statistik uji *Breusch-Pagan* sebesar 16,066 dan *P-value* sebesar 0,006659. Dengan taraf signifikansi 5% sehingga diperoleh keputusan tolak H_0 yang artinya variansi antar lokasi berbeda atau terdapat perbedaan karakteristik antara satu titik pengamatan dengan titik pengamatan lainnya.

Berdasarkan pengujian aspek data spasial yaitu heterogenitas spasial diketahui bahwa data jumlah kasus diabetes melitus di Jawa Timur memenuhi aspek data spasial, sehingga dapat dilanjutkan pemodelan dengan menggunakan metode GWGPR dan GWNBR.

I. Pemodelan GWGPR Jumlah Kasus Diabetes Melitus

Model GWGPR menggunakan pembobot geografis pada penaksiran parameternya. Pengujian parameter pada model GWGPR terdiri dari pengujian secara serentak dan parsial. Pengujian secara serentak bertujuan untuk mengetahui apakah variabel prediktor memberikan pengaruh yang signifikan terhadap model. Pengujian secara serentak untuk model GWGPR menggunakan hipotesis sebagai berikut.

$$H_0 : \beta_1(u_i, v_i) = \dots = \beta_5(u_i, v_i) = 0 ; i = 1,2, \dots, 38$$

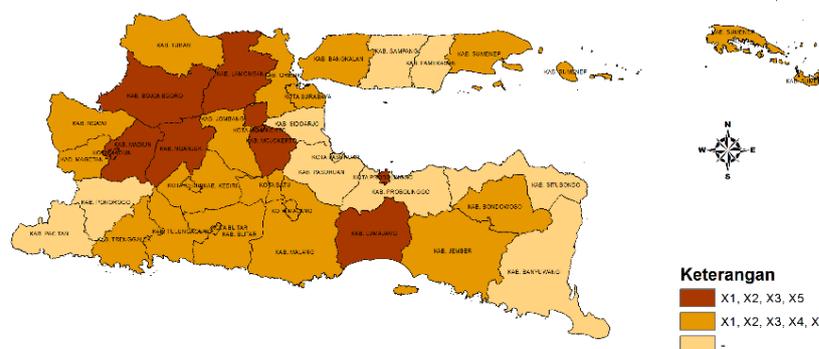
$$H_1 : \text{minimal ada satu } \beta_j(u_i, v_i) \neq 0 ; j = 1,2,3,4,5$$

Nilai devians model GWGPR yang didapatkan adalah sebesar 5012969. Taraf signifikansi yang digunakan yaitu 5% ($\alpha = 5\%$), sehingga didapatkan nilai $\chi^2_{(0,05;5)}$ sebesar 11,0705. Hal ini menunjukkan bahwa nilai devians lebih besar dari nilai $\chi^2_{(0,05;5)}$, sehingga diperoleh keputusan tolak H_0 yang artinya minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap model regresi GWGPR. Kemudian dilanjutkan dengan pengujian secara parsial untuk mengetahui variabel yang berpengaruh signifikan pada masing-masing kabupaten/kota di Jawa Timur dengan hipotesis sebagai berikut.

$$H_0 : \beta_1(u_i, v_i) = 0$$

$$H_1 : \beta_j(u_i, v_i) \neq 0$$

Hasil pengujian secara parsial menghasilkan parameter yang berbeda untuk setiap kabupaten/kota di Jawa Timur. Statistik uji dapat dilihat berdasarkan nilai $|Z_{hitung}|$ masing-masing parameter yang dibandingkan dengan $Z_{(0,05/2)}$ adalah 1,96. Apabila nilai $|Z_{hitung}| > 1,96$ maka keputusan adalah tolak H_0 , yang artinya parameter signifikan terhadap model. Hasil pengelompokan dengan menggunakan metode GWGPR ditunjukkan pada gambar sebagai berikut.



Gambar 3 Pengelompokan Kabupaten/Kota di Jawa Barat Berdasarkan Variabel yang Signifikan dengan Metode GWGPR

Pada Gambar 3 di atas diketahui terdapat 3 kategori yaitu yang pertama adalah daerah yang signifikan untuk semua variabel prediktor yaitu terdapat 19 Kabupaten/Kota, kedua adalah daerah yang signifikan oleh empat prediktor yaitu terdapat di 9 Kabupaten/Kota, dan ketiga adalah daerah yang tidak signifikan oleh semua prediktor yaitu terdapat 10 Kabupaten/Kota. Setelah didapatkan variabel yang signifikan untuk setiap daerah, selanjutnya berdasarkan pengujian parameter secara parsial akan disajikan sebagai contoh pembahasan hasil pengujian tersebut pada lokasi penelitian yang sama dengan contoh pemodelan metode GWGPR, menggunakan lokasi ke-8 (u_8, v_8) yaitu Kabupaten Lumajang. Hasil estimasi parameter untuk Kabupaten Lumajang dapat ditunjukkan pada Tabel 8.

Tabel 8 Nilai Estimasi Parameter Model GWGPR di Kabupaten Lumajang

Parameter	Estimasi	SE($\hat{\beta}$)	Z hitung	P-value
$\hat{\beta}_0$	0,013009	0,00004407	295,1811	0,0000000*
$\hat{\beta}_1$	0,137500	0,00074326	184,9953	0,0000000*
$\hat{\beta}_2$	0,223032	0,00101904	218,8652	0,0000000*
$\hat{\beta}_3$	0,362427	0,00349412	103,7248	0,0000000*
$\hat{\beta}_4$	0,000283	0,00068673	0,412098	0,6802675
$\hat{\beta}_5$	0,482782	0,00371754	129,8660	0,0000000*
$\hat{\theta}$	0,015257			
Devians	5012969			

*) signifikan dengan taraf 5%

Berdasarkan Tabel 8 dapat ditunjukkan nilai $\hat{\theta}$ yang lebih besar dari nol yang artinya terjadi kasus overdispersi. Dari hasil analisis didapatkan nilai devians sebesar 5012969. Taraf signifikansi yaitu 5%, sehingga nilai $\chi^2_{(0,05;5)}$ sebesar 11,0705. Hal ini menunjukkan nilai devians lebih besar dari $\chi^2_{(0,05;5)}$ sehingga diperoleh keputusan tolak H_0 yang artinya minimal ada satu variabel prediktor yang signifikan terhadap model GWGPR di Kabupaten Lumajang.

Pengujian parameter secara parsial dilakukan menggunakan statistik uji $|Z_{hitung}|$ yang kemudian dibandingkan dengan $Z_{(\alpha/2)}$. Tabel 8 menunjukkan bahwa terdapat empat variabel prediktor yang memiliki nilai $|Z_{hitung}| > 1,96$ yaitu rasio tingkat pendidikan SMA ke-atas, proporsi aktivitas fisik, persentase konsumsi rokok, dan proporsi konsumsi minuman manis, maka didapatkan keputusan tolak H_0 , yang artinya parameter signifikan terhadap model regresi di Kabupaten Lumajang. Persamaan model GWGPR di Kabupaten Lumajang adalah sebagai berikut.

$$\hat{\mu} = \exp(0,0130090 + 0,1375007X_1 + 0,223032X_2 + 0,362427X_3 + 0,000283X_4 + 0,482782X_5)$$

Berdasarkan variabel yang signifikan dari model yang terbentuk di Kabupaten Lumajang dapat disimpulkan bahwa setiap kenaikan satu satuan rasio tingkat pendidikan SMA ke-atas (X_1), maka akan meningkatkan rata-rata jumlah kasus diabetes melitus sebesar $\exp(0,137500) = 1,147402$ kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan. Hal ini bisa jadi karena adanya faktor lain yang dapat meningkatkan jumlah kasus diabetes melitus. Setiap kenaikan satu satuan proporsi aktivitas fisik penduduk (X_2) maka akan meningkatkan rata-rata jumlah kasus diabetes melitus sebesar $\exp(0,223032) = 1,249861$ kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan. Hal ini berarti bahwa semakin banyak kegiatan aktivitas fisik yang dilakukan oleh masyarakat maka akan meningkatkan kemungkinan terkena kasus diabetes melitus. Hal ini bisa jadi karena adanya faktor lain seperti riwayat penyakit yang dapat meningkatkan jumlah kasus diabetes melitus. Setiap kenaikan satu satuan persentase konsumsi rokok oleh penduduk (X_3) maka akan meningkatkan rata-rata jumlah kasus diabetes melitus sebesar $\exp(0,362427) = 1,436812$ kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan. Hal ini menandakan bahwa semakin banyak masyarakat mengonsumsi rokok maka memiliki kemungkinan lebih besar terkena kasus diabetes melitus. Setiap kenaikan satu satuan proporsi konsumsi minuman manis (X_5) maka akan meningkatkan rata-rata kasus diabetes melitus sebesar $\exp(0,482782) = 1,620577$ kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan. Hal ini telah sesuai bahwa semakin banyak orang mengonsumsi minuman manis maka akan meningkatkan kemungkinan terkena penyakit diabetes melitus.

J. Pemodelan GWNBR Jumlah Kasus Diabetes Melitus

Model GWNBR menggunakan pembobot geografis pada penaksiran parameternya. Pengujian parameter pada model GWNBR terdiri dari pengujian secara serentak dan parsial. Pengujian secara serentak bertujuan untuk mengetahui

apakah variabel prediktor memberikan pengaruh yang signifikan terhadap model. Pengujian secara serentak untuk model GWNBR menggunakan hipotesis sebagai berikut.

$$H_0 : \beta_1(u_i, v_i) = \dots = \beta_5(u_i, v_i) = 0; i = 1, 2, \dots, 38$$

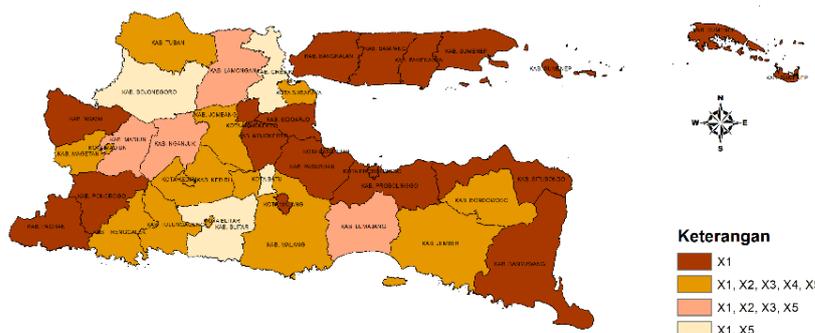
$$H_1 : \text{minimal ada satu } \beta_j(u_i, v_i) \neq 0; j = 1, 2, 3, 4, 5$$

Nilai devians model GWNBR yang didapatkan adalah sebesar 3060638. Taraf signifikansi yang digunakan yaitu 5% ($\alpha = 5\%$), sehingga didapatkan nilai $\chi^2_{(0,05;5)}$ sebesar 11,0705. Hal ini menunjukkan bahwa nilai devians lebih besar dari nilai $\chi^2_{(0,05;5)}$, sehingga diperoleh keputusan tolak H_0 yang artinya minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap model regresi GWNBR. Kemudian dilanjutkan dengan pengujian secara parsial untuk mengetahui variabel yang berpengaruh Signifikan pada masing-masing kabupaten/kota di Jawa Timur dengan hipotesis sebagai berikut.

$$H_0 : \beta_1(u_i, v_i) = 0$$

$$H_1 : \beta_j(u_i, v_i) \neq 0$$

Hasil pengujian secara parsial menghasilkan parameter yang berbeda untuk setiap kabupaten/kota di Jawa Timur. Statistik uji dapat dilihat berdasarkan nilai $|Z_{hitung}|$ masing-masing parameter yang dibandingkan dengan $Z_{(0,05/2)}$ adalah 1,96. Apabila nilai $|Z_{hitung}| > 1,96$ maka keputusan adalah tolak H_0 , yang artinya parameter signifikan terhadap model. Hasil pengelompokan dengan menggunakan metode GWNBR ditunjukkan pada gambar sebagai berikut.



Gambar 4 Pengelompokan Kabupaten/Kota di Jawa Barat Berdasarkan Variabel yang Signifikan dengan Metode GWNBR

Berdasarkan Gambar 4 di atas diketahui bahwa terdapat empat kelompok Kabupaten/Kota berdasarkan variabel yang signifikan. Setelah didapatkan variabel yang signifikan untuk setiap daerah, selanjutnya berdasarkan pengujian parameter secara parsial akan disajikan sebagai contoh pembahasan hasil pengujian tersebut pada lokasi penelitian yang sama dengan contoh pemodelan metode GWNBR, menggunakan lokasi ke-8 (u_8, v_8) yaitu Kabupaten Lumajang. Hasil estimasi parameter untuk Kabupaten Lumajang dapat ditunjukkan pada Tabel 9.

Tabel 9 Nilai Estimasi Parameter Model GWNBR di Kabupaten Lumajang

Parameter	Estimasi	SE($\hat{\beta}$)	Z hitung	P-value
$\hat{\beta}_0$	0,020702	0,00005523	374,8141	0,000000*
$\hat{\beta}_1$	0,219298	0,00093184	235,3382	0,000000*
$\hat{\beta}_2$	0,353737	0,00127751	276,8954	0,000000*
$\hat{\beta}_3$	0,577304	0,00438044	131,7913	0,000000*
$\hat{\beta}_4$	0,000771	0,00085959	0,896936	0,369753
$\hat{\beta}_5$	0,768356	0,00466020	164,8762	0,000000*
$\hat{\theta}$	0,019066			
Devians	3060638			

*) signifikan dengan taraf 5%

Berdasarkan Tabel 9 dapat ditunjukkan nilai $\hat{\theta}$ yang lebih besar dari nol yang artinya terjadi kasus overdispersi. Dari hasil analisis didapatkan nilai devians sebesar 3060638. Taraf signifikansi yaitu 5%, sehingga nilai $\chi^2_{(0,05;5)}$ sebesar 11,0705. Hal ini menunjukkan nilai devians lebih besar dari $\chi^2_{(0,05;5)}$ sehingga diperoleh keputusan tolak H_0 yang artinya minimal ada satu variabel prediktor yang signifikan terhadap model regresi GWNBR di Kabupaten Lumajang.

Pengujian parameter secara parsial dilakukan menggunakan statistik uji $|Z_{hitung}|$ yang kemudian dibandingkan dengan $Z_{(\alpha/2)}$. Tabel 9 menunjukkan bahwa terdapat empat variabel prediktor yang memiliki nilai $|Z_{hitung}| > 1,96$ yaitu rasio tingkat pendidikan SMA ke-atas, proporsi aktivitas fisik, persentase konsumsi rokok, dan proporsi konsumsi minuman manis, maka didapatkan keputusan tolak H_0 , yang artinya parameter signifikan terhadap model regresi di Kabupaten Lumajang. Persamaan model GWNBR di Kabupaten Lumajang adalah sebagai berikut.

$$\hat{\mu} = \exp(0,020702 + 0,219298X_1 + 0,353737X_2 + 0,577304X_3 + 0,000771X_4 + 0,768356X_5)$$

Berdasarkan variabel yang signifikan dari model yang terbentuk di Kabupaten Lumajang dapat disimpulkan bahwa setiap kenaikan satu satuan rasio tingkat pendidikan SMA ke-atas (X_1), maka akan meningkatkan rata-rata jumlah kasus diabetes melitus sebesar $\exp(0,219298) = 1,245202$ kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan. Hal ini bisa jadi karena adanya faktor lain yang dapat meningkatkan jumlah kasus diabetes melitus. Setiap kenaikan satu satuan proporsi aktivitas fisik penduduk (X_2) maka akan meningkatkan rata-rata jumlah kasus diabetes melitus sebesar

$\exp(0,353737) = 1,424381$ kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan. Hal ini bisa jadi karena adanya faktor lain seperti riwayat penyakit yang dapat meningkatkan jumlah kasus diabetes melitus. Setiap kenaikan persentase konsumsi rokok yang dihisap penduduk (X_3) maka akan meningkatkan rata-rata jumlah kasus diabetes melitus sebesar $\exp(0,577304) = 1,78123$ kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan. Hal ini menandakan bahwa semakin banyak masyarakat mengonsumsi rokok maka memiliki kemungkinan lebih besar terkena kasus diabetes melitus. Setiap kenaikan satu satuan proporsi konsumsi minuman manis (X_5) maka akan meningkatkan rata-rata kasus diabetes melitus sebesar $\exp(0,768356) = 2,156219$ kali lipat dari nilai sebelumnya dengan asumsi variabel lain konstan. Hal ini telah sesuai bahwa semakin banyak orang mengonsumsi minuman manis maka akan meningkatkan kemungkinan terkena penyakit diabetes melitus.

K. Pemilihan Model Terbaik

Pemilihan model terbaik didasarkan pada nilai AICc. Model yang baik adalah model yang memiliki nilai AICc yang kecil. Nilai AICc di setiap model regresi spasial disajikan pada Tabel 10.

Metode	AICc
GWGPR	2653,440
GWNBR	4248,041

Pada Tabel 10 dapat dilihat bahwa hasil AICc GWGPR dan GWNBR jauh berbeda. Dengan demikian model GWGPR lebih sesuai untuk memodelkan faktor-faktor yang mempengaruhi jumlah kasus diabetes melitus di setiap Kabupaten/Kota di Provinsi Jawa Timur karena memiliki nilai AICc lebih kecil daripada model GWNBR. Hal ini karena pada model GWGPR terdapat Kabupaten/Kota yang tidak signifikan untuk semua variabel prediktor dan ada pula Kabupaten/Kota yang signifikan terhadap semua variabel prediktor, sehingga terdapat perbedaan antar wilayah yang cukup signifikan.

V. ANALISIS DAN PEMBAHASAN

A. Kesimpulan

Berdasarkan hasil analisis dan pembahasan maka dapat disimpulkan beberapa hal sebagai berikut.

1. Pada tahun 2019 jumlah kasus diabetes melitus di Provinsi Jawa Timur mencapai 841.994, dengan rata-rata sebesar 22.258 kasus per kabupaten/kota. Kota Surabaya memiliki jumlah kasus diabetes melitus paling banyak yaitu sejumlah 94.076 kasus, sedangkan Kota Batu memiliki jumlah kasus diabetes melitus paling sedikit yaitu 3.344 kasus. Kasus diabetes melitus memiliki standar deviasi sebesar 18.273 yang menandakan persebaran data yang tinggi, hal itu terjadi karena terdapat perbedaan signifikan antara jumlah kasus diabetes melitus pada setiap kabupaten/kota.
2. Hasil pemodelan dengan metode GWGPR menghasilkan tiga kelompok dan GWNBR menghasilkan empat kelompok daerah menurut variabel yang berpengaruh signifikan terhadap jumlah kasus diabetes melitus di kabupaten/kota di Jawa Timur. Hasil pemodelan GWGPR dan GWNBR menghasilkan faktor-faktor yang signifikan berpengaruh terhadap jumlah kasus diabetes melitus di kabupaten/kota di Jawa Timur adalah rasio tingkat pendidikan SMA ke-atas, proporsi aktivitas fisik penduduk, persentase konsumsi rokok tembakau, persentase penderita hipertensi, dan proporsi konsumsi minuman manis.
3. Kriteria AICc menunjukkan bahwa metode GWGPR merupakan metode yang paling sesuai untuk memodelkan jumlah kasus diabetes melitus di Kabupaten/Kota di Provinsi Jawa Timur dibandingkan metode GWNBR karena nilai AICc yang lebih kecil. Hal ini karena pada model GWGPR terdapat Kabupaten/Kota yang tidak signifikan untuk semua variabel prediktor dan ada pula Kabupaten/Kota yang signifikan terhadap semua variabel prediktor, sehingga terdapat perbedaan antar wilayah yang cukup signifikan.

B. Saran

Dari analisis yang telah dilakukan didapatkan beberapa variabel yang berpengaruh signifikan terhadap jumlah kasus diabetes melitus untuk setiap kabupaten/kota di Jawa Timur. berdasarkan model GWGPR dan GWNBR didapatkan pengelompokan menurut variabel yang signifikan di tiap kabupaten/kota sehingga diharapkan kedepannya ada pengurangan rata-rata jumlah kasus diabetes melitus dengan cara mengimplementasikan pola hidup berdasarkan variabel yang berpengaruh signifikan di tiap lokasi. Untuk penelitian kedepannya perlu adanya penambahan variabel prediktor yang memberikan pengaruh meningkatnya jumlah kasus diabetes melitus di Provinsi Jawa Timur. Selain itu, apabila terdapat variabel bebas yang memberikan pengaruh signifikan pada semua lokasi maka pada penelitian selanjutnya dapat melakukan analisis dengan metode *Mixed Geographically Weighted Generalized Poisson Regression* atau *Mixed Geographically Weighted Negative Binomial Regression*.

DAFTAR PUSTAKA

- [1] Kementerian Kesehatan RI, Infodatin Diabetes Melitus 2020, Jakarta: Kemenkes RI, 2020.
- [2] Dinas Kesehatan Provinsi Jawa Timur, Profil Kesehatan Jawa Timur 2020, Surabaya: Dinas Kesehatan Jawa Timur, 2020.
- [3] Quinones, S., Goyal, A & Ahmed, Z.U., "Geographically Weighted Machine Learning Model for Untangling Spatial Heterogeneity of Type 2 Diabetes Melitus (T2D) Prevalence in the USA," *Scientific Reports*, pp. 1-13, 2021.
- [4] Kauh, B., Schweikart, J., Krafft, T., & Andrea, "Do The Risk Factors for Type 2 Diabetes Mellitus Vary by Location? A Spatial Analysis of Health Insurance Claims Northeastern Germany Using Kernel Density Estimation and Geographically Weighted Regression," *International Journal of Health Geographics*, 2016.
- [5] A. Fotheringham, C. Brunson, & M. C. T. Nayaka, "Geographically Weighted Poisson Regression for Disease Association Mapping," *Statistics in Medicine*, pp. 2695-2717, 2005.
- [6] R. Walpole, Pengantar Metode Statistika, B. Sumantri, Penyunt., Jakarta: PT Gramedia Pusaka Utama, 1995.
- [7] Fatma, Desya, "Peta Tematik : Pengertian, Ciri-ciri, Jenis dan Contohnya," 6 Oktober 2017. [Online]. Available: <https://ilmugeografi.com/kartografi/peta-tematik>.
- [8] Gujarati & Porter, Dasar-Dasar Ekonometrika Edisi 5, Jakarta: Salemba Empat, 2010.
- [9] A. M. Law, & W. D. Kelton, Simulation Modeling & Analysis, New York: McGraw Hill, Inc, 2000.
- [10] Trivedi, A., Cameron, C., & Pravin, K, "Regression-Based Test for Overdispersion in The Poisson Model," *Journal of Econometrics*, vol. 46, pp. 347-364, 1990.
- [11] Fotheringham, S. A., Brunson, C., & Charlton, M., Geographically Weighted Regression : the analysis of spatially varying relationship, West Sussex: John Wiley & Sons Ltd, 2002.
- [12] Johnson, R. A. & Wichern, D. W., Applied Multivariate Statistical Analysis Third Edition, New Jersey, 1992.
- [13] F. Famoye, "On The Generalizes Poisson Regression Model with an Application to Accident Data," *Journal of Data Science*, pp. 287-295, 2004.
- [14] Fitri, E. U. L., Pemodelan Faktor-Faktor yang Mempengaruhi Jumlah Kasus Tuberkulosis di Jawa Timur Menggunakan Metode Geographically Weighted Generalized Poisson Regression dan Geographically Weighted Negative Binomial Regression, Surabaya: Institut Teknologi Sepuluh Nopember, 2017.
- [15] F. N. Nayaka, "Geographically Weighted Poisson Regression for Disease Association Mapping," *Statistics in Medicine*, pp. 2695-2717, 2005.
- [16] Yasin, R.E & Cakra, H., Geographically Weighted Regression (GWR) Sebuah Pendekatan Regresi Geografis, Yogyakarta: MOBIUS, 2017.
- [17] W. Greene, Functional Form and Heterogeneity in Models for Count Data, New York: New York University, 2007.
- [18] Hilbe, J. M., Negative Binomial Regression, New York : Cambridge University Press, 2011.
- [19] A. R. d. Silva, & T. C. Rodrigues, "Geographically Weighted Negative Binomial Regression Incorporating Overdispersion," *Statistics and Computing*, 2013.
- [20] Burnham, K. P. & Anderson, D. R., Model Selection and Multimodel Inference : A Practical Information-Theoretic Approach, Springer: Verlag New York, Inc, 2002.
- [21] PERKENI, Konsensus Pengelolaan Diabetes Melitus di Indonesia, Jakarta: PB. PERKENI, 2011.
- [22] WHO, Global Report on Diabetes, Switzerland: WHO, 2016.
- [23] Kementerian Kesehatan RI, Laporan Riskesdas Jawa Timur 2018, Jakarta: Lembaga Penerbit Badan Litbang Kesehatan, 2018.
- [24] M. Indah, Pusat Data dan Informasi Kementerian Kesehatan RI : Tuberkulosis, Jakarta: Kementerian Kesehatan RI, 2018.
- [25] Humas Dinas Kesehatan Provinsi Jawa Timur, Profil Kesehatan Provinsi Jawa Timur, Surabaya: Dinas Kesehatan Provinsi Jawa Timur, 2020.
- [26] D. Gujarati, Basic Econometrics (Ekonometrika Dasar). (S. Zain, Penerj.), Jakarta: Erlangga, 2013.
- [27] A. Agresti, An Introduction to Categorical Data Analysis, 2 penyunt., Hoboken: John Wiley & Sons, Inc, 2019.
- [28] Trivedi, A., Cameron, C., & Pravin, K, Regression Analysis of Count Data, USA: Cambridge University Press, 2013.
- [29] F. Famoye, "On The Generalizes Poisson Regression Model with an Application to Accident Data," *Journal of Data Science*, pp. 287-295, 2004.



© 2023 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (<http://creativecommons.org/licenses/by-sa/4.0/>).