

Penerapan Model Regresi Ensemble Non-Hybrid pada Data Kemiskinan di Provinsi Jawa Tengah

Cornelia Ardiana Savita¹, Sri Sulistijowati Handajani², Bowo Winarno³

^{1,3} Program Studi Matematika FMIPA, Universitas Sebelas Maret

² Program Studi Statistika FMIPA, Universitas Sebelas Maret

Email: corneliardiana@student.uns.ac.id

Abstrak

Keywords:
Kemiskinan, regresi spasial, ensemble non-hybrid

Kemiskinan merupakan salah satu permasalahan pada negara berkembang seperti Indonesia. Provinsi Jawa Tengah memiliki jumlah penduduk miskin yang cukup tinggi dan berada di posisi 13 dari 33 Provinsi di Indonesia. Kemiskinan di suatu wilayah dipengaruhi kedekatan wilayah di sekelilingnya sehingga digunakan model regresi spasial. Tujuan penelitian ini adalah menerapkan model regresi spasial ensemble pada data kemiskinan dengan pembobot persinggungan sisi-sudut di Provinsi Jawa Tengah. Model regresi spasial ensemble digunakan karena dapat mengurangi keragaman serta memberikan peningkatan akurasi pada model prediksi. Hasil penelitian menunjukkan bahwa model regresi spasial yang digunakan adalah model regresi spasial eror. Karena hanya efek ketergantungan spasial eror yang memenuhi sehingga teknik ensemble yang digunakan adalah teknik ensemble non-hybrid. Berdasarkan hasil penelitian diperoleh model regresi spasial eror ensemble untuk persentase penduduk miskin di Provinsi Jawa Tengah dengan faktor yang mempengaruhi adalah rumah tangga yang menggunakan jamban sendiri, rumah tangga yang pernah membeli beras miskin, dan laju pertumbuhan penduduk.

1. PENDAHULUAN

Kemiskinan merupakan permasalahan mendasar dalam pembangunan ekonomi yang sedang dihadapi hampir setiap negara. Sebagai negara berkembang, Indonesia memiliki tingkat kemiskinan yang cukup tinggi dibandingkan beberapa negara di sekitarnya. Menurut *World Bank* dalam Saputra [7], salah satu penyebab kemiskinan karena kurangnya pendapatan dan aset untuk memenuhi kebutuhan dasar seperti makanan, pakaian, perumahan, tingkat kesehatan, dan pendidikan yang seharusnya diperoleh masyarakat.

Provinsi Jawa Tengah memiliki luas wilayah 32.543 km² atau sekitar 28,94% dari

luas pulau Jawa, serta jumlah penduduknya mencapai 33,77 juta jiwa. Tingkat kemiskinan terbesar kedua di pulau Jawa adalah Provinsi Jawa Tengah, yaitu sebesar 14,51%. Badan Pusat Statistik [2] mencatat jumlah penduduk miskin di Provinsi Jawa Tengah pada tahun 2014 sebesar 4,56 juta jiwa dan mengalami peningkatan menjadi 4,57 juta jiwa pada tahun 2015.

Nasir dalam Prasetyo [6] menyatakan bahwa permasalahan kemiskinan merupakan permasalahan yang kompleks dan bersifat multidimensional. Oleh karena itu, upaya yang harus dilakukan untuk mengatasi masalah kemiskinan adalah dengan mengidentifikasi faktor-faktor yang

berpengaruh terhadap kemiskinan. Kemiskinan di suatu daerah tidak lepas dari pengaruh kemiskinan di wilayah sekelilingnya. Hal ini mengindikasikan adanya pengaruh spasial. Hukum pertama geografi yang menjadi salah satu dasar pengembangan analisis spasial dikemukakan oleh Tobler yang menyatakan segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang lebih dekat akan lebih berpengaruh daripada sesuatu yang lebih jauh (Anselin [1]).

LeSage [4] menyatakan jika informasi ruang atau spasial diabaikan maka pengamatan menghasilkan kesimpulan yang berbeda sehingga model yang terbentuk menjadi tidak layak. Cara mengatasi permasalahan tersebut dengan memasukkan efek spasial antar wilayah ke dalam model, sehingga model yang digunakan adalah model regresi spasial. Model yang digunakan adalah model spasial lag dan model spasial eror yang didasarkan pada efek ketergantungan spasial lag dan spasial eror menggunakan pendekatan area.

Tingkat akurasi model regresi spasial dapat ditingkatkan dengan tujuan memberikan hasil yang lebih baik serta dapat menurunkan tingkat kesalahan dengan melakukan metode resampling dalam penyusunan modelnya. Milton et al. [5] menyatakan, teknik ensemble dapat digunakan untuk mengurangi keragaman yang terdapat pada model prediksi, serta terbukti dapat memberikan peningkatan akurasi prediksi. Prinsip teknik ensemble adalah memberikan keakuratan prediksi yang lebih baik dengan menggabungkan hasil estimasi parameter beberapa model menjadi satu estimasi parameter. Canuto et al. [3] menyatakan terdapat dua teknik ensemble, yaitu teknik ensemble hybrid dan teknik ensemble non-hybrid. Teknik ensemble hybrid adalah teknik yang menggunakan berbagai jenis model, sedangkan teknik ensemble non-hybrid adalah teknik yang hanya menggunakan satu jenis model. Model

regresi spasial ensemble yang digunakan dalam penelitian ini adalah model regresi spasial ensemble non-hybrid. Berdasarkan uraian tersebut dalam penelitian ini akan diterapkan model regresi spasial ensemble non-hybrid pada data kemiskinan di Provinsi Jawa Tengah.

2. METODE

Data yang digunakan pada penelitian ini adalah data kemiskinan di Provinsi Jawa Tengah tahun 2015 yang diperoleh dari Badan Pusat Statistik (BPS [2]). Faktor-faktor yang mempengaruhi kemiskinan dapat diidentifikasi dengan persentase penduduk miskin Y sebagai variabel dependen. Sedangkan variabel independen yang digunakan yaitu persentase penduduk miskin usia 15 tahun ke atas yang tidak tamat SD , persentase angka melek huruf usia 15-55 tahun , persentase penduduk miskin usia 15 tahun ke atas dan tidak bekerja , persentase penduduk miskin usia 15 tahun ke atas yang bekerja di sektor pertanian , persentase perempuan pengguna alat kontrasepsi , persentase perempuan berstatus miskin usia 15-49 tahun yang persalinan pertamanya ditolong oleh tenaga kesehatan , persentase rumah yang memiliki luas lantai perkapita kurang dari 8 m² , persentase rumah tangga yang menggunakan jamban sendiri , persentase rumah tangga yang pernah membeli beras raskin , dan persentase laju pertumbuhan penduduk .

Berikut adalah langkah-langkah yang dilakukan untuk mencapai tujuan penelitian.

Menentukan model regresi linear menggunakan metode stepwise dan menguji asumsi pada model regresi linear.

Menentukan matriks pembobot spasial W persinggungan Queen terstandardisasi.

Menguji adanya autokorelasi spasial menggunakan Moran's I.

Menguji efek ketergantungan spasial menggunakan uji Lagrange Multiplier (LM).

Melakukan pendugaan dan pengujian parameter model regresi spasial SAR atau SEM.

Melakukan prediksi model regresi spasial dengan teknik ensemble non-hybrid dengan langkah

membangkitkan noise dengan

$$z \sim N(0, \sigma),$$

menambahkan noise pada data persentase penduduk miskin (Y),

melakukan pengujian pada data yang telah diberi additive noise dengan model regresi SAR atau SEM,

mengulangi langkah (a)-(c) sebanyak p kali dengan additive noise yang berbeda pada tiap iterasinya,

membuat prediksi model regresi spasial ensemble non-hybrid dengan menghitung rata-rata hasil estimasi parameter p model regresi spasial.

menguji asumsi pada model regresi spasial ensemble yang dihasilkan dan mengukur kebaikan model.

3. HASIL DAN PEMBAHASAN

3.1 Model Regresi Linear

Pembentukan model regresi linear diawali dengan melakukan seleksi variabel yang signifikan terhadap model. Seleksi variabel dilakukan dengan metode *stepwise* antara persentase penduduk miskin dengan variabel yang mempengaruhinya. Pada kasus kemiskinan di Provinsi Jawa Tengah tahun 2015, model regresi linear yang diperoleh yaitu

$$\hat{Y} = 12,9293 - 0,0732X_8 + 0,1289X_9 - 13,6754X_{10}$$

dengan (X_8) adalah persentase rumah tangga yang menggunakan jamban sendiri, (X_9) adalah persentase rumah tangga yang pernah membeli beras miskin, (X_{10}) adalah

persentase laju pertumbuhan penduduk. Uji asumsi kenormalan, homoskedastisitas, dan non multikolinearitas dilakukan untuk mengetahui apakah model regresi yang telah didapatkan memenuhi asumsi regresi.

a. Uji kenormalan

Uji kenormalan dilakukan menggunakan uji Kolmogorov-Smirnov. Hipotesis H_0 : residu berdistribusi normal dan H_1 : residu tidak berdistribusi normal.

DK uji ini adalah $\{D_{hitung} | D_{hitung} > D_{\alpha,n}\}$ dengan

$$D_{\alpha,n} = D_{0,05;35} = 0,224. \quad H_0 \text{ ditolak}$$

apabila nilai $D_{hitung} \in DK$. Diperoleh nilai D_{hitung} sebesar

$$\max |F_0(X_i) - S_n(X_i)| = 0,0729.$$

Karena nilai $D_{hitung} \notin DK$ sehingga H_0 tidak ditolak yang berarti error mengikuti distribusi normal.

b. Uji non multikolinearitas

Salah satu cara untuk mengetahui ada atau tidak adanya multikolinearitas dapat dilihat berdasarkan nilai *Variance Inflation Factors* (VIF) yang kurang dari

$$10 \text{ dengan } VIF_k = \frac{1}{1 - R_k^2}. \text{ Nilai VIF}$$

$X_8 = 1,0785$, $X_9 = 1,0989$, dan $X_{10} = 1,0599$ kurang dari 10 yang sehingga disimpulkan bahwa tidak terdapat multikolinearitas.

c. Uji homoskedastisitas

Uji homoskedastisitas dilakukan menggunakan uji Breusch Pagan (BP).

Hipotesis H_0 : tidak terdapat heteroskedastisitas dan H_1 : terdapat heteroskedastisitas. DK uji ini adalah

$$\{BP | BP > \chi_{\alpha,k}^2\} \text{ dengan}$$

$$\chi_{(0,05;3)}^2 = 7,815. \quad H_0 \text{ ditolak apabila}$$

nilai $BP \in DK$. Diperoleh nilai $BP = 7,8908$. Karena nilai $BP \in DK$

sehingga H_0 ditolak yang berarti terdapat heteroskedastisitas.

3.2 Indeks Moran's I

Indeks Moran's I digunakan untuk mengukur hubungan spasial atau autokorelasi antara pengamatan pada suatu lokasi dengan lokasi lain yang berdekatan. Indeks Moran's I dalam model regresi spasial ditentukan menggunakan nilai error model regresi. Nilai Indeks Moran I dapat diperoleh sebagai berikut

$$IM = \frac{n \sum_{i=1}^n \sum_{j \neq i}^n w_{ij} (\varepsilon_i - \bar{\varepsilon})(\varepsilon_j - \bar{\varepsilon})}{\left(\sum_{i=1}^n \sum_{j \neq i}^n w_{ij} \right) \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2}$$

Berdasarkan perhitungan, nilai IM sebesar 0,24 sehingga menunjukkan adanya autokorelasi spasial positif yang berarti adanya kemiripan nilai error dari lokasi-lokasi yang berdekatan dan nilai error tersebut cenderung berkelompok.

3.3 Uji Efek Spasial

Dalam menentukan model regresi spasial yang akan digunakan dilakukan pengujian efek spasial. Efek spasial yaitu ketergantungan spasial yang terjadi karena adanya korelasi antar wilayah. Efek ketergantungan spasial yang digunakan yaitu ketergantungan lag dan error yang diuji menggunakan uji Lagrange *Multiplier* (LM).

a. Uji Lagrange *Multiplier* (LM) lag

Uji LM lag digunakan untuk mengetahui adanya efek spasial dalam variabel dependen dengan $H_0 : \rho = 0$ (tidak terdapat efek spasial lag) dan $H_1 : \rho \neq 0$ (terdapat efek spasial lag). Daerah kritis untuk uji ini adalah $\{LM_\rho | LM_\rho > \chi^2_{(0,05;1)} = 3,851\}$ dengan H_0 ditolak jika $LM_\rho \in DK$. Diperoleh

$$\text{nilai } LM_\rho = \left(\frac{\boldsymbol{\varepsilon}' \mathbf{W} \mathbf{y}}{\boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} n^{-1}} \right)^2 \frac{1}{\mathbf{H}} = 2,1087.$$

Karena nilai $LM_\rho \notin DK$ sehingga H_0 tidak ditolak yang berarti tidak terdapat efek spasial lag dalam model.

b. Uji Lagrange *Multiplier* (LM) error

Uji LM error digunakan mengetahui adanya efek spasial dalam error dengan $H_0 : \lambda = 0$ (tidak terdapat efek spasial error) dan $H_1 : \lambda \neq 0$ (terdapat efek spasial error). Daerah kritis untuk uji ini adalah $\{LM_\lambda | LM_\lambda > \chi^2_{(0,05;1)} = 3,851\}$ dengan H_0 ditolak jika $LM_\lambda \in DK$.

Diperoleh nilai

$$LM_\lambda = \left(\frac{\boldsymbol{\varepsilon}' \mathbf{W} \mathbf{y}}{\boldsymbol{\varepsilon}' \boldsymbol{\varepsilon} n^{-1}} \right)^2 \frac{1}{\text{tr}(\mathbf{W}' \mathbf{W} + \mathbf{W}^2)}$$

$$= 4,0997.$$

Karena nilai $LM_\lambda \in DK$ sehingga H_0 ditolak yang berarti terdapat efek spasial error dalam model.

3.4 Model Regresi Spasial

Menurut Anselin [1], model regresi spasial merupakan suatu model regresi yang mengandung efek spasial, yaitu autokorelasi spasial dalam bentuk matriks adalah

$$\mathbf{Y} = \rho \mathbf{W} \mathbf{Y} + \mathbf{X} \boldsymbol{\beta} + \mathbf{u}, \mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \boldsymbol{\varepsilon}$$

dengan \mathbf{Y} adalah vektor variabel dependen, ρ adalah koefisien parameter spasial *autoregressive*, \mathbf{W} adalah matriks pembobot, \mathbf{X} adalah matriks variabel independen, $\boldsymbol{\beta}$ adalah vektor parameter regresi, \mathbf{u} adalah vektor error, λ adalah koefisien parameter spasial error, $\boldsymbol{\varepsilon}$ adalah vektor error berdistribusi normal.

Berdasarkan uji LM didapatkan bahwa model regresi spasial yang tepat untuk data persentase kemiskinan di Provinsi Jawa Tengah adalah model regresi spasial error. Hasil estimasi parameter pada model spasial

eror dilakukan uji signifikansi dengan $H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ (parameter tidak signifikan) dan H_1 : paling tidak ada satu $\beta_i \neq 0, i = 1, \dots, 4$ yang signifikan. DK uji ini adalah $\{ |z_{hitung}| > z_{\alpha/2} = 1,96 \}$. H_0 ditolak apabila $|z_{hitung}| \in DK$. Hasil estimasi parameter model spasial eror terdapat pada Tabel 1.

Tabel 3.1. Nilai estimasi parameter dan z_{hitung} model spasial eror

Variabel	Nilai estimasi	z_{hitung}
Konstanta	11,9264	3,9519
λ	0,4270	2,3217
X_8	-0,0668	-2,1015
X_9	0,1204	6,2223
X_{10}	-10,9425	-1,9885

Berdasarkan Tabel 3.1. Diperoleh model spasial eror yang dinyatakan sebagai

$$\hat{Y} = 11,9264 - 0,0668X_8 + 0,1204X_9 - 10,9425X_{10} + \mathbf{u},$$

$$\mathbf{u} = 0,4270\mathbf{W}\mathbf{u}$$

Variabel rumah tangga yang menggunakan jamban sendiri (X_8), rumah tangga yang pernah membeli beras miskin (X_9), dan laju pertumbuhan penduduk (X_{10}) berpengaruh signifikan pada model karena $|z_{hitung}| \in DK$. Model tersebut memiliki nilai R^2 sebesar 0,7045. Hal ini menunjukkan 70,45% persentase penduduk miskin di Provinsi Jawa Tengah tahun 2015 dapat dijelaskan oleh variabel independen tersebut.

3.5 Model Regresi Spasial Ensemble

Model regresi spasial *ensemble* adalah model yang dihasilkan dari rata-rata hasil estimasi parameter p model regresi, dengan p adalah banyaknya model regresi spasial. Model regresi spasial *ensemble* dinyatakan sebagai

$$\hat{Y} = \frac{1}{Q} \sum_{p=1}^Q \hat{Y}_p$$

dengan \hat{Y} adalah rata-rata estimasi parameter, Q adalah banyaknya resampling, dan \hat{Y}_p adalah model regresi spasial ke- p .

Model regresi spasial *ensemble* yang digunakan adalah model regresi spasial *ensemble non-hybrid* karena hanya efek ketergantungan spasial eror yang terpenuhi sehingga model yang dibentuk adalah model regresi spasial eror. Pada data persentase penduduk miskin akan ditambahkan *noise* (z) yang berasal dari bangkitan data $\mathbf{z} \sim N(0, \sigma)$ yang dituliskan sebagai

$$\mathbf{Y} + \mathbf{z} = \mathbf{Y}_p$$

dengan \mathbf{Y} adalah vektor variabel dependen sebelum ditambahkan *noise* dan \mathbf{Y}_p adalah vektor variabel dependen setelah ditambahkan *noise*.

Nilai σ yang digunakan adalah 2,53. Penambahan *noise* dilakukan sebanyak 100 kali sehingga didapatkan 100 model regresi spasial eror sebagai berikut

$$\hat{Y}_1 = 8,0574 - 0,0358X_8 + 0,1679X_9 - 16,7820X_{10} + \mathbf{u}, \mathbf{u} = 0,2608\mathbf{W}\mathbf{u}$$

$$\hat{Y}_2 = 16,7957 - 0,1128X_8 + 0,1090X_9 - 14,1988X_{10} + \mathbf{u}, \mathbf{u} = 0,4517\mathbf{W}\mathbf{u}$$

$$\hat{Y}_3 = 8,1945 - 0,0416X_8 + 0,1849X_9 - 17,9414X_{10} + \mathbf{u}, \mathbf{u} = 0,2806\mathbf{W}\mathbf{u}$$

$$\begin{aligned} & \vdots \\ \hat{Y}_{100} &= 13,7713 - 0,0733X_8 + 0,1192X_9 \\ & - 15,9774X_{10} + \mathbf{u}, \mathbf{u} = 0,0773\mathbf{Wu} \end{aligned}$$

Model regresi spasial eror *ensemble* diperoleh dengan menghitung rata-rata hasil estimasi parameter dari 100 model sehingga hanya terdapat satu model tunggal regresi spasial eror. Model regresi spasial eror *ensemble* dinyatakan sebagai

$$\begin{aligned} \hat{Y} &= 12,6847 - 0,0743X_8 + 0,1325X_9 \\ & - 12,9863X_{10} + \mathbf{u}, \\ \mathbf{u} &= 0,2226\mathbf{Wu} \end{aligned}$$

Model tersebut memiliki nilai R^2 sebesar 0,7359. Hal ini menunjukkan 73,59% persentase penduduk miskin di Provinsi Jawa Tengah tahun 2015 dapat dijelaskan oleh variabel independen tersebut. Selanjutnya dilakukan uji kenormalan dan homoskedastisitas untuk mengetahui apakah model memenuhi asumsi regresi. Uji kenormalan dilakukan menggunakan uji Kolmogorov-Smirnov. Diperoleh hasil nilai $D = 0,075 < D_{0,05;35} = 0,224$ yang berarti eror mengikuti distribusi normal. Uji homoskedastisitas dilakukan menggunakan uji Breusch-Pagan (BP). Diperoleh hasil nilai $BP = 7,339 < \chi^2_{0,05;3} = 7,815$ yang berarti tidak terdapat heteroskedastisitas.

Berdasarkan model regresi spasial eror *ensemble* untuk persentase penduduk miskin, dapat diinterpretasikan bahwa koefisien λ menunjukkan jika suatu wilayah dikelilingi oleh beberapa wilayah lain, maka pengaruh wilayah-wilayah yang mengelilinginya dapat diukur sebesar 0,2226 dikalikan dengan rata-rata eror spasial disekitarnya. Sehingga meningkatkan nilai persentase penduduk miskin akibat dari pengaruh eror disekitar wilayah tersebut. Koefisien variabel rumah tangga yang menggunakan jamban bersama

dan laju pertumbuhan penduduk bernilai negative. Hal ini menunjukkan untuk setiap kenaikan rumah tangga yang menggunakan jamban sendiri (X_8) dan laju pertumbuhan penduduk (X_{10}) sebesar satu satuan akan menurunkan persentase penduduk miskin (Y) sebesar 0,0743% dan 12,9863%. Koefisien rumah tangga yang pernah membeli beras miskin bernilai positif. Hal ini menunjukkan untuk setiap kenaikan rumah tangga yang pernah membeli beras miskin (X_9) sebesar satu satuan akan meningkatkan persentase penduduk miskin (Y) sebesar 0,1325%.

4 KESIMPULAN

Berdasarkan hasil dan pembahasan dapat diperoleh kesimpulan bahwa faktor yang berpengaruh terhadap kemiskinan di Provinsi Jawa Tengah tahun 2015 adalah persentase rumah tangga yang menggunakan jamban sendiri (X_8), persentase rumah tangga yang pernah membeli beras miskin (X_9), dan persentase laju pertumbuhan penduduk (X_{10}). Model regresi spasial *ensemble non-hybrid* terbukti dapat mengurangi keragaman dan meningkatkan prediksi dengan R^2 sebesar 73,59%. Model regresi eror *ensemble* dinyatakan sebagai

$$\begin{aligned} \hat{Y} &= 12,6847 - 0,0743X_8 + 0,1325X_9 \\ & - 12,9863X_{10} + \mathbf{u}, \\ \mathbf{u} &= 0,2226\mathbf{Wu} \end{aligned}$$

REFERENSI

- [1] Anselin L., *Spatial econometrics: Method and model. Sixth Edition.* Dordrecht: Kluwer Academic Publishers; 1998. 1-978.
- [2] Badan Pusat Statistik. *Data dan informasi kemiskinan kabupaten/kota.* Indonesia: BPS; 2015.

- [3] Canuto AMP, Abreu MCC, Oliveira L, Junior JCX, Santos A. Investigating the influence of the choice of the ensemble members in accuracy and diversity of selection-based and fusion-based methods for ensemble. *Pattern Recognition Letters*. 2007; 28:472-486.
- [4] LeSage JP. Bayesian estimation of spatial autoregressive models. *International Regional Science Review*. 1997;(1-2):113-129.
- [5] Milton JN, Steinberg MH, Sebastiani P.. Evaluation of an ensemble of genetic models for prediction of a quantitative trait. *Methods Article*. 2015; 5:1-6.
- [6] Prasetyo A. Analisis Faktor-Faktor yang Mempengaruhi Tingkat Kemiskinan. Universitas Diponegoro:2010.
- [7] Saputra WA. Analisis Pengaruh Jumlah Penduduk, PDRB, IPM, Pengangguran terhadap Tingkat Kemiskinan di Kabupaten/Kota Jawa Tengah. Universitas Diponegoro; 2011.

