

## INTISARI

### **ANALISIS *VARIABLE IMPORTANCE* TERHADAP VARIABEL RESPON DALAM PEMODELAN REGRESI SPASIAL MENGGUNAKAN LMG DAN *RANDOM FOREST***

Oleh

RAZIEQ ILHAM AMALIY  
20/466541/PPA/06107

Penelitian ini menggunakan dua metode pengukuran *variable importance* untuk mengidentifikasi kontribusi dari faktor-faktor dalam menjelaskan distribusi spasial kemiskinan, yaitu metode Lindeman, Merenda, dan Gold (LMG) dalam regresi linear berganda dan *Random Forest* dalam *machine learning*. Dalam penelitian ini, kasus kemiskinan yang terjadi di Pulau Sumatera dalam tahun 2021 dianalisis menggunakan model spasial. Penelitian ini menggunakan model *Spatial Lag of X* (SLX) untuk memodelkan kemiskinan. Model ini mempertimbangkan adanya pengaruh faktor-faktor kemiskinan di lokasi tetangga terhadap kemiskinan di suatu lokasi. Selanjutnya, metode *stepwise selection* digunakan untuk mengeliminasi variabel *redundant* agar mendapatkan model SLX yang terbaik. Variabel-variabel dalam model ini selanjutnya dianalisis menggunakan LMG dan RF dengan tujuan mendapatkan faktor utama kemiskinan. Berdasarkan hasil analisis data, LMG dan RF konsisten dalam dua faktor utama kemiskinan, yaitu rata-rata Pengeluaran Per Kapita Disesuaikan lokasi tetangga (W\_PPK) dan Pengeluaran Per Kapita Disesuaikan suatu lokasi (PPK). Berdasarkan metode LMG, rata-rata Pengeluaran Per Kapita Disesuaikan lokasi tetangga (W\_PPK) dan Pengeluaran Per Kapita Disesuaikan memiliki kepentingan berturut-turut 24,023% dan 16,485% dalam memodelkan kemiskinan suatu lokasi. Sedangkan, Pengeluaran Per Kapita Disesuaikan lokasi tetangga (W\_PPK) dan Pengeluaran Per Kapita Disesuaikan memiliki kepentingan berturut-turut 16,404% dan 13,338% dalam memodelkan kemiskinan suatu lokasi menurut metode *random forest*.

**Kata Kunci:** *Variable importance*, regresi spasial, *LMG*, *random forest*

## ABSTRACT

### ***VARIABLE IMPORTANCE ANALYSIS OF RESPONSE VARIABLE IN SPATIAL REGRESSION USING LMG AND RANDOM FOREST***

By

RAZIEQ ILHAM AMALIY

20/466541/PPA/06107

*This study uses two methods of variable importance measures to identify the contribution of factors in explaining the spatial distribution of poverty, namely the Lindeman, Merenda, and Gold (LMG) method in multiple linear regression and Random Forest in machine learning. In this research, poverty cases that occurred on Sumatera Island in 2021 were analyzed using a spatial model. This research uses the Spatial Lag of X (SLX) model to explain the poverty. This model considers the influence of poverty factors in neighboring locations on poverty in a location. Next, the stepwise selection method is used to eliminate redundant variables in order to get the best SLX model. The variables in this model are then analyzed using LMG and RF with the aim of obtaining the main factors of poverty. Based on the results of data analysis, LMG and RF are consistent in two main factors of poverty, namely average Per Capita Expenditure Adjusted for neighboring locations (W\_PPK) and Per Capita Expenditure Adjusted for a location (PPK). Based on the LMG method, the average Adjusted Per Capita Expenditure of neighboring locations (W\_PPK) and Adjusted Per Capita Expenditure have respectively 24.023% and 16.485% importance in modeling the poverty of a location. Meanwhile, Adjusted Per Capita Expenditure for neighboring locations (W\_PPK) and Adjusted Per Capita Expenditure have respectively 16.404% and 13.338% importance in modeling the poverty of a location according to the random forest method.*

**Keywords:** *Variable importance, spatial regression, LMG, random forest*