



## Metode *Least Absolute Shrinkage And Selection Operator* (Lasso) Untuk Penanganan Masalah Multikolinearitas (Studi Kasus: Jumlah Kejahatan Yang Dilaporkan (*Crime Total*) Di Provinsi Sulawesi Tenggara Tahun 2022)

### INFO PENULIS

Irma Yahya  
Universitas Halu Oleo  
[IrmaYahya@gmail.com](mailto:IrmaYahya@gmail.com)

Indah Sahara  
Universitas Halu Oleo  
[indahsaharaaa@gmail.com](mailto:indahsaharaaa@gmail.com)

Muhammad Ihwal  
Universitas Halu Oleo  
[MuhammadIhwal@gmail.com](mailto:MuhammadIhwal@gmail.com)

Makkulau  
Universitas Halu Oleo  
[Makkulau@gmail.com](mailto:Makkulau@gmail.com)

### INFO ARTIKEL

ISSN: 3026-3603  
Vol. 4, No. 1 April 2026  
<http://jurnal.ardenjaya.com/index.php/ajst>

© 2026 Arden Jaya Publisher All rights reserved

### **Saran Penulisan Referensi:**

Baharuddin., Nugiansa, S., Ihwal, M., Makkulau., Laome, L., & Ampa, A. T. (2026) Pemodelan Spasial Indeks Pembangunan Manusia di Pulau Sulawesi Menggunakan Spatial Error Model. *Arus Jurnal Sains dan Teknologi*, 4 (1), 14-25.

### **Abstrak**

Jumlah kejahatan yang dilaporkan di Provinsi Sulawesi Tenggara pada tahun 2022 mengalami peningkatan, sehingga perlu dianalisis faktor-faktor yang memengaruhinya. Namun, penerapan regresi linear berganda sering menghadapi masalah multikolinearitas antarvariabel independen yang dapat mengurangi keakuratan model. Penelitian ini bertujuan mengatasi multikolinearitas dengan menerapkan metode *Least Absolute Shrinkage and Selection Operator* (LASSO). Data yang digunakan berupa data sekunder dari Badan Pusat Statistik dengan delapan variabel independen. Metode LASSO diterapkan dengan algoritma LARS dan Cross Validation (CV) 5-fold untuk menentukan parameter *shrinkage* optimum. Lambda optimum adalah parameter regulasi yang fungsinya untuk mengontrol seberapa besar penalti yang diberikan terhadap koefisien regresi yang diperoleh dari nilai CV MSE yang paling minimum. Hasil analisis menunjukkan metode LASSO berhasil mengatasi multikolinearitas dengan nilai  $R^2$  sebesar 86,66% dan Variation Inflation Factor (VIF) variabel Rata-rata Lama Sekolah, Tingkat Pengangguran Terbuka, Kepadatan Penduduk, Gini Rasio, Jumlah Penduduk, dan Rasio Jenis Kelamin, masing-masing sebesar 4,74; 2,43; 4,68; 1,58; 1,48; dan 1,63, dalam model akhir lebih kecil dari 5. Berdasarkan nilai  $R^2$  dan nilai VIF menunjukkan metode LASSO dapat menjadi metode yang efektif untuk membangun model regresi yang stabil pada kasus dengan multikolinearitas tinggi.

**Kata Kunci:** CV 5-fold, Kejahatan, LASSO, Multikolinearitas, Regresi, VIF

## Abstract

The number of reported crimes in Southeast Sulawesi Province increased in 2022, thus it is necessary to analyze the influencing factors. However, the application of multiple linear regression often encounters multicollinearity problems among the independent variables, which can reduce the accuracy of the model. This study aims to address multicollinearity by applying the Least Absolute Shrinkage and Selection Operator (LASSO) method. The data used are secondary data from the Central Bureau of Statistics with eight independent variables. The LASSO method was applied using the LARS algorithm and 5-fold Cross Validation (CV) to determine the optimum shrinkage parameter. The optimum lambda is a regularization parameter that controls the amount of penalty applied to the regression coefficients, determined by the minimum CV MSE value. The analysis results show that the LASSO method successfully addresses multicollinearity, with an  $R^2$  value of 86.66% and the Variation Inflation Factor (VIF) values of the variables Average Years of Schooling, Open Unemployment Rate, Population Density, Gini Ratio, Total Population, and Sex Ratio being 4.74; 2.43; 4.68; 1.58; 1.48; and 1.63, respectively — all below 5 in the final model. Based on the  $R^2$  and VIF values, the LASSO method proves to be an effective alternative for building a stable regression model in cases with high multicollinearity.

**Keywords:** 5-fold CV, Crime, LASSO, Multicollinearity, Regression, VIF

## A. Pendahuluan

Jumlah kejahatan di Provinsi Sulawesi Tenggara pada tahun 2019, 2020 hingga 2021 masing-masing tercatat sebanyak 4.098, 3.901 dan 3.000 kasus kejahatan (BPS, 2022). Data ini menunjukkan adanya penurunan jumlah kasus kejahatan yang cukup signifikan. Namun pada tahun 2022, Badan Pusat Statistik (BPS) Provinsi Sulawesi Tenggara mencatat peningkatan jumlah kasus kejahatan menjadi 3.561 kasus.

Analisis regresi merupakan salah satu metode dalam statistika yang banyak digunakan untuk memodelkan, menganalisis kekuatan hubungan dan memprediksi pengaruh antara satu atau beberapa variabel independen terhadap variabel dependen. Salah satu asumsi yang harus dipenuhi dalam regresi linear berganda adalah tidak terjadinya multikolinearitas (Lestari *et al.*, 2022).

Multikolinearitas dapat menyebabkan kesulitan dalam interpretasi koefisien regresi dan meningkatkan varians dari prediksi yang dihasilkan, sehingga berpotensi mengarah pada kesimpulan yang tidak tepat. Menurut Lestari *et al.* (2022).

Salah satu metode yang terbukti efektif dalam mengatasi multikolinearitas adalah *Least Absolute Shrinkage and Selection Operator* (LASSO). Metode LASSO merupakan teknik regularisasi yang tidak hanya mengurangi varians estimasi tetapi juga melakukan seleksi variabel dengan menjadikan beberapa koefisien regresi menjadi nol (0). Metode LASSO dapat menghasilkan model yang lebih *robust* dengan melakukan penalti pada ukuran koefisien (Hastie *et al.*, 2015). Penelitian ini berfokus pada penggunaan metode LASSO untuk menangani multikolinearitas dalam analisis jumlah kejahatan di Provinsi Sulawesi Tenggara.

## Tinjauan Pustaka

### 1. Analisis Regresi Berganda

Istilah regresi pertama kali diperkenalkan oleh seorang ahli yang bernama Fancis Galton pada tahun 1886. Menurutnya, analisis regresi berkenaan dengan studi ketergantungan dari suatu variabel yang disebut dengan variabel tak independen (respon) terhadap variabel independen (prediktor).

Analisis regresi linear berganda menjadi analisis diantara dua atau lebih variabel independen  $(X_1, X_2, X_3, \dots, X_n)$  dengan variabel dependen (Y) dalam melihat tingkat hubungannya secara linear. Menurut Sulantari *et al.*, (2022), secara sistematis model dari regresi linear berganda adalah:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_n \quad (1)$$

dimana  $Y$  adalah variabel dependen,  $\beta_0$  adalah *intercept*/titik potong antara sumbu tegak  $y$  dan garis fungsi linear,  $\beta_1, \dots, \beta_n$  adalah koefisien regresi dan  $X_1, \dots, X_p$  adalah variabel independen ( $p = 1, 2, 3, \dots, n$ ).

## 2. Penduga Parameter Regresi Berganda

Salah satu metode estimasi parameter untuk regresi linear berganda adalah Metode Kuadrat Terkecil (MKT). Pendugaan parameter model regresi linear berganda untuk populasi tidak dapat diestimasi secara langsung, sehingga dapat diestimasi dengan data sampel dalam bentuk sebagai berikut.

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_p \quad (2)$$

## 3. Multikolinearitas

Multikolinearitas adalah suatu kondisi dimana terjadi korelasi antara variabel independen atau antar variabel independen tidak bersifat saling bebas. Istilah multikolinearitas atau kolinearitas ganda diciptakan oleh Ragner Frish yang menyatakan multikolinearitas terjadi jika adanya hubungan linear sempurna diantara variabel-variabel independen dalam model regresi (Supriyadi *et al.*, 2017).

Menurut (Wasailane, *et al.*, 2014) hal-hal utama yang menyebabkan terjadinya multikolinearitas dalam model regresi antara lain: (1) kesalahan teoritis dalam pembentukan model fungsi regresi yang dipergunakan dan (2) terlampau kecilnya jumlah pengamatan yang akan dianalisis dengan model regresi. Ada beberapa cara yang dapat digunakan untuk mendeteksi adanya multikolinearitas, diantaranya adalah sebagai berikut.

## 4. Variance Inflation Factor (VIF)

VIF digunakan sebagai kriteria untuk mendeteksi multikolinearitas pada regresi linier yang melibatkan lebih dari dua variabel bebas. Dikutip dari Nurdin *et al.*, (2018), nilai VIF yang lebih besar dari 5 sudah sangat mengkhawatirkan. VIF untuk koefisien regresi- $j$  diidentifikasi sebagai berikut.

$$VIF_{(j)} = \frac{1}{(1 - R_j^2)} \quad j = 1, 2, 3, \dots, n \quad (3)$$

$R_j^2$  merupakan koefisien determinasi yang didapat dari variabel independen  $X_j$  yang diregresikan dengan variabel independen lainnya. Jika nilai  $VIF_{(j)}$  lebih besar dari 5 maka terjadi masalah multikolinearitas.

## 5. Korelasi Pearson

Salah satu perhitungan koefisien korelasi adalah dengan korelasi pearson yang digunakan untuk melihat hubungan diantara variabel (Pendi, 2021). Korelasi pearson pertama kali diperkenalkan oleh Karl Pearson pada tahun 1990 (Firdaus & Zamal, 2009). Nilai korelasi pearson antara  $x$  dan  $y$  dapat dihitung dengan rumus sebagai berikut.

$$r = \frac{n \left( \sum_{i=1}^n xy \right) - \left( \sum_{i=1}^n x \right) \left( \sum_{i=1}^n y \right)}{\sqrt{\left\{ n \sum_{i=1}^n x^2 - \left( \sum_{i=1}^n x \right)^2 \right\} \left\{ n \sum_{i=1}^n y^2 - \left( \sum_{i=1}^n y \right)^2 \right\}}} \quad (4)$$

Korelasi dilambangkan dengan nilai  $r$ , dengan ketentuan  $-1 \leq r \leq 1$ . Apabila  $r = -1$  artinya negatif sempurna,  $r = 0$  tidak ada korelasi, dan  $r = 1$  artinya korelasi sangat kuat. Untuk mengetahui korelasi antara dua variabel, hipotesis yang digunakan adalah sebagai berikut.

$H_0 : \rho = 0$  (tidak terdapat korelasi)

$H_1 : \rho \neq 0$  (terdapat korelasi)

## 6. Least Absolute Shrinkage and Selection Operator (LASSO)

LASSO merupakan suatu metode yang diperkenalkan oleh Tibshirani pada tahun 1996. Sesuai dengan namanya, regresi LASSO merupakan metode regresi berganda yang digunakan untuk *shrinkage* yaitu menyusutkan koefisien taksiran hingga mendekati angka nol dan *selection operator* yaitu menyeleksi variabel-variabel independen sehingga menghasilkan model dengan variabel terbaik. Regresi LASSO juga digunakan untuk data yang kontinu dan memerlukan variabel independen yang berdistribusi normal baku (Ningsih *et al.*, 2023). Regresi LASSO didefinisikan sebagai berikut.

$$\hat{\beta}_{LASSO} = \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 \quad (5)$$

dengan fungsi kendala  $\sum_{j=1}^p |\beta_j| \leq t$ .

Solusi koefisien lasso diperoleh dengan menentukan parameter tuning yang dibakukan, yaitu  $S = \frac{t}{\sum_{j=1}^p |\beta_j^0|}$  dengan  $t = \sum_{j=1}^p |\beta_j|$ . Pada saat  $s = 1$ , maka diperoleh penduga kuadrat terkecil.

Semakin nilai  $s$  mendekati nol, maka koefisien LASSO disusutkan ke nol. Sehingga  $s$  memiliki rentang nilai 0 sampai dengan 1. Parameter tuning ( $s$ ) ini berikutnya disebut dengan istilah parameter penyusutan atau *shrinkage* parameter. Dikutip dari Hastie *et al.*, (2008), metode LASSO secara umum dinyatakan sebagai berikut.

$$Y_{LASSO} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i + \varepsilon \quad (6)$$

## 7. Algoritma LARS

Untuk menghasilkan parameter regresi ( $\beta_0, \beta_1, \dots, \beta_j$ ) pada Persamaan (6) tidak dapat dilakukan dengan proses diferensiasi seperti halnya metode OLS, karena persamaan tersebut merupakan bentuk yang tidak dapat diturunkan pada semua titik. *Least Angle Regression* (LARS) merupakan suatu metode regresi yang algoritmanya dapat dimodifikasi menjadi algoritma komputasi untuk metode LASSO. Dengan memodifikasi algoritma LARS dapat memberikan koefisien taksiran metode LASSO. Algoritma LARS memberikan jalan yang efisien dalam menyelesaikan regresi LASSO (Robbani *et al.*, 2019).

## 8. Validasi Silang (Cross Validation)

Validasi silang atau *Cross Validation* (CV) adalah salah satu teknik yang digunakan dalam memperkirakan galat prediksi untuk meningkatkan keakuratan pemilihan model (James *et al.*, 2013). Validasi silang digunakan untuk memprediksi model dan memperkirakan seberapa akurat sebuah model.

Menurut Efron & Tibshirani (1993), salah satu pendekatan dari CV yaitu dengan menerapkan *k-fold* di mana data dibagi ke dalam *k-fold*. Dengan *fold* pertama sebagai data untuk validasi dan *k-1 fold* sisanya untuk melatih model. Proses akan terus berulang hingga mencapai *fold* terakhir yang digunakan sebagai data validasi. Metode ini baik digunakan ketika jumlah data amatan sedikit. Formula *k-fold* CV dituliskan sebagai berikut.

$$CV = \frac{1}{k} \sum_{i=1}^k \left( y_i - y_{-i(k)} \right)^2 \quad (7)$$

dimana  $y_{-i(k)}$  adalah dugaan  $y$  saat *fold* ke- $k$  tidak digunakan dalam mengestimasi model. Dalam *k-fold* CV sering kali menggunakan  $k = 5$  atau  $k = 10$ . Dikutip dari James *et al.*, (2013), keuntungan dalam menggunakan *5-fold* atau *10-fold* yaitu akan menghasilkan ragam rendah. Dengan *k-fold* CV, maka diperoleh tuning parameter optimum dilihat dari nilai CV yang terkecil.

## 9. Evaluasi Performa Model

Evaluasi performa model bertujuan untuk mengukur seberapa baik model dapat memprediksi data baru. Perhitungan metrik kinerja seperti *Mean Squared Error* (MSE) dan *R-squared* ( $R^2$ ) membantu mengevaluasi seberapa baik model LASSO melakukan prediksi. Meskipun model LASSO menggunakan regularisasi dan penalti, evaluasi kinerja tetap penting untuk memastikan bahwa model memprediksi dengan akurat.

### a. Mean Squared Error (MSE)

*Mean Squared Error* (MSE) adalah metrik yang mengukur rata-rata selisih kuadrat antara nilai sebenarnya dengan nilai yang diprediksi oleh model. MSE dapat dihitung dengan formula sebagai berikut.

$$MSE = \frac{1}{2} \sum_{i=1}^n \left( y_i - \hat{y}_i \right)^2 \quad (8)$$

Semakin kecil MSE, maka semakin baik pula model dalam memprediksi, dengan selisih kecil antara nilai prediksi dan nilai sebenarnya.

### b. Koefisien Determinasi ( $R^2$ )

Koefisien determinasi atau *R-squared* digunakan untuk melihat seberapa besar variabel independen dapat dijelaskan oleh variabel dependen (Fatimah, 2022). Nilai  $R^2$  berkisar antara 0 hingga 1, dimana semakin nilai  $R^2$  mendekati 1 berarti model semakin baik dalam menjelaskan data. Dikutip dari Gujarati (2013), perhitungan pada  $R^2$  dapat didefinisikan sebagai berikut.

$$R^2 = \frac{\sum_{i=1}^n (y - \bar{y})^2}{\sum_{i=1}^n (y - \bar{y})^2} = \frac{JKR}{JKT} = 1 - \frac{JKG}{JKT} \quad (9)$$

Semakin tinggi nilai  $R^2$  yang dihasilkan suatu model, maka semakin baik pula variabel-variabel independen dalam model tersebut dalam menjelaskan variabel dependen. Sifat dari koefisien determinasi yakni memiliki nilai yang tak negatif dan terletak pada interval  $0 \leq R^2 \leq 1$ .

## 10. Jumlah Kejahatan yang Dilaporkan

Kejahatan meliputi segala tingkah laku manusia walaupun tidak ditentukan oleh undang-undang, tetapi oleh warga masyarakat dirasakan atau ditafsirkan sebagai tingkah laku atau perbuatan yang secara ekonomis atau psikologis menyerang dan melukai perasaan susila dalam kehidupan bersama (Muliadi, 2012). Faktor-faktor yang berkaitan dengan jumlah kejahatan yang dilaporkan (*crime total*) adalah sebagai berikut.

### a. Jumlah Penduduk Miskin

Penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita perbulan dibawah garis kemiskinan. Kemiskinan dapat menjadikan efek yang cukup serius bagi pembangunan manusia karena masalah kemiskinan merupakan sebuah masalah yang kompleks yang sebenarnya bermula dari kemampuan daya beli masyarakat yang tidak mampu untuk mencukupi kebutuhan pokok sehingga kebutuhan yang lain terabaikan.

### b. Rata-rata Lama Sekolah

Rata-rata lama sekolah (RLS) adalah angka yang menggambarkan lamanya (tahun) masa sekolah yang dialami penduduk usia 25 tahun ke atas, dimana perhitungannya didasarkan pada asumsi bahwa pada usia 25 tahun, mayoritas penduduk telah menyelesaikan pendidikan formalnya. Semakin tinggi tingkat pendidikan maka semakin besar pula kemungkinan seseorang mendapatkan upah yang lebih tinggi. Hal ini menciptakan harapan bahwa peluang seseorang untuk terlibat dalam aktivitas Kriminal/kejahatan akan berkurang (Priatna, 2015).

### c. Tingkat Pengangguran Terbuka

Tingkat Pengangguran Terbuka merupakan persentase pengangguran terbuka terhadap angkatan kerja. Pengangguran terbuka merupakan salah satu jenis pengangguran dimana angkatan kerja tidak memiliki pekerjaan sama sekali.

### d. Kepadatan Penduduk

Kepadatan penduduk adalah perbandingan antara jumlah penduduk dengan luas wilayah yang dihuni. Jumlah penduduk yang semakin banyak dan tidak terdistribusi dengan merata akan menimbulkan konflik (Desinta, 2022). Semakin meningkat jumlah penduduk maka peluang terjadinya tindak kejahatan juga semakin meningkat.

### e. Gini Rasio

Rasio Gini adalah ukuran yang menunjukkan tingkat ketimpangan pengeluaran secara keseluruhan, dengan nilai 0 hingga 1. Nilai yang lebih tinggi menunjukkan tingkat ketimpangan yang lebih besar (BPS, 2023).

### f. Rata-rata Upah/Gaji Bersih Sebulan Pekerja Formal

Rata-rata Upah/Gaji Bersih Sebulan Pekerja Formal adalah jumlah rata-rata gaji atau penghasilan yang diterima oleh pekerja yang bekerja di sektor formal setelah dipotong pajak, iuran sosial, dan potongan lainnya dalam satu bulan (BPS, 2022).

### g. Jumlah Penduduk

Penduduk adalah semua orang yang berdomisili di wilayah geografis Republik Indonesia selama enam bulan atau lebih dan mereka yang berdomisili kurang dari enam bulan tetapi bertujuan untuk menetap (Arum & Haris, 2019).

### h. Rasio Jenis Kelamin

Perbandingan jumlah laki-laki dan perempuan di suatu tempat dan pada waktu tertentu disebut rasio jenis kelamin, yang biasanya digambarkan sebagai banyaknya laki-laki per 100 perempuan (BPS, 2023).

## B. Metodologi

Materi penelitian yang digunakan adalah berupa data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Sulawesi Tenggara pada situs <https://sultra.bps.go.id/>. Data yang digunakan merupakan data jumlah kejahatan yang dilaporkan (*crime total*) berdasarkan Kabupaten/Kota di Provinsi Sulawesi Tenggara pada tahun 2022 beserta faktor-faktor yang diduga mempengaruhinya.

### Tahapan Analisis Data

Langkah-langkah analisis data dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan variabel penelitian
2. Mendeteksi adanya multikolinieritas dengan memperhatikan korelasi antarvariabel independen dan nilai *variance inflation factor* (VIF).
3. Memutuskan hasil pendeteksian adanya multikolinieritas
4. Langkah mengatasi multikolinieritas dengan menggunakan metode *Least Absolute Shrinkage and Selection Operator* (LASSO) dengan algoritma LARS dengan langkah-langkah sebagai berikut.
  - a. Melakukan standarisasi data sehingga memiliki nilai tengah nol dan ragam satu.
  - b. Menentukan nilai  $\lambda$  optimum dengan validasi silang lipat-5 (*5-fold cross validation*)
  - c. Melakukan pemodelan regresi LASSO berdasarkan nilai  $\lambda$  optimum
  - d. Evaluasi performa model
  - e. Interpretasi model

## C. Hasil dan Pembahasan

### Deskripsi Data

Analisis deskriptif data jumlah kejahatan yang dilaporkan (*crime total*) di Provinsi Sulawesi Tenggara beserta faktor-faktor yang mempengaruhinya pada tahun 2022 yang masing-masing berjumlah 17 data yang diambil berdasarkan kabupaten/kota ditunjukkan pada Tabel 4.1 berikut.

**Tabel 1** Analisis Deskriptif Data

Variabel	Mean	Standar Deviasi	Maksimum	Minimum
Y	150,353	191,083	778	0
X <sub>1</sub>	18,223	9,140	35,79	5,47
X <sub>2</sub>	8,866	1,283	12,52	7,33
X <sub>3</sub>	3,048	1,128	5,39	1,47
X <sub>4</sub>	189,471	323,690	1342	17
X <sub>5</sub>	0,348	0,036	0,29	0,43
X <sub>6</sub>	2731807	592858,4	3952152	1905399
X <sub>7</sub>	158921,2	91678,71	38383	356747
X <sub>8</sub>	102,338	3,091	96,16	106,97

Berdasarkan Tabel 4.1 diketahui bahwa jumlah penduduk miskin (X<sub>1</sub>) paling tinggi di Provinsi Sulawesi Tenggara pada tahun 2022 adalah 35,79 dan yang paling rendah adalah 5,47 dengan rata-rata jumlah penduduk miskin (X<sub>1</sub>) sebesar 18,223 dan standar deviasi sebesar 9,140. Selanjutnya diketahui bahwa rata-rata lama sekolah (X<sub>2</sub>) paling tinggi di Provinsi Sulawesi Tenggara adalah 12,52 dan yang paling rendah adalah 7,33 dengan rata-rata sebesar 8,866 dan standar standar deviasi sebesar 1,283. Deskripsi data lainnya juga secara jelas dirincikan pada Tabel 4.1.

### Mendeteksi Multikolinieritas

#### 1. Menghitung Nilai Korelasi antar Variabel Independen

Mendeteksi multikolinieritas melalui nilai korelasi adalah dengan cara menghitung korelasi dari setiap pasang variabel independen hingga membentuk matriks korelasi berukuran 8x8. Nilai korelasi yang mendekati 1 atau *p-value* kurang dari 0,05 mengindikasikan adanya multikolinieritas pada pasangan variabel tersebut.

Nilai korelasi linear antara variabel independen dapat dihitung menggunakan Persamaan 2.14. Nilai matriks korelasi Pearson disajikan dalam Tabel 2 berikut.

**Tabel 2** Nilai Korelasi Variabel Independen

		X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
X <sub>2</sub>	Korelasi	-0,034						
	<i>p-value</i>	0,898						
X <sub>3</sub>	Korelasi	0,155	0,570					
	<i>p-value</i>	0,553	0,017					
X <sub>4</sub>	Korelasi	-0,053	0,792	0,733				
	<i>p-value</i>	0,839	0,000	0,000				
X <sub>5</sub>	Korelasi	-0,111	0,581	0,256	0,459			
	<i>p-value</i>	0,671	0,014	0,320	0,063			
X <sub>6</sub>	Korelasi	0,524	0,624	0,272	0,423	0,327		
	<i>p-value</i>	0,031	0,007	0,290	0,090	0,200		
X <sub>7</sub>	Korelasi	0,805	0,447	0,505	0,521	0,207	0,757	
	<i>p-value</i>	0,000	0,072	0,039	0,032	0,426	0,000	
X <sub>8</sub>	Korelasi	0,169	0,082	-0,292	-0,278	-0,066	0,444	0,068
	<i>p-value</i>	0,517	0,754	0,255	0,280	0,800	0,074	0,796

Berdasarkan Tabel 4.2, diketahui beberapa variabel yang saling berkorelasi diantaranya adalah variabel jumlah penduduk miskin (X<sub>1</sub>) dengan rata-rata upah/gaji bersih sebulan pekerja formal (X<sub>6</sub>), jumlah penduduk miskin (X<sub>1</sub>) dengan jumlah penduduk (X<sub>7</sub>), variabel rata-rata lama sekolah (X<sub>2</sub>) dengan tingkat pengangguran terbuka (X<sub>3</sub>), variabel rata-rata lama sekolah (X<sub>2</sub>) dengan kepadatan penduduk (X<sub>4</sub>), variabel rata-rata lama sekolah (X<sub>2</sub>) dengan gini rasio (X<sub>5</sub>), variabel rata-rata lama sekolah (X<sub>2</sub>) dengan rata-rata upah/gaji bersih sebulan pekerja formal (X<sub>6</sub>), variabel tingkat pengangguran terbuka (X<sub>3</sub>) dengan kepadatan penduduk (X<sub>4</sub>), variabel tingkat pengangguran terbuka (X<sub>3</sub>) dengan jumlah penduduk (X<sub>7</sub>), variabel kepadatan penduduk (X<sub>4</sub>) dengan jumlah penduduk (X<sub>7</sub>), dan variabel rata-rata upah/gaji bersih sebulan pekerja formal (X<sub>6</sub>) dengan jumlah penduduk (X<sub>7</sub>). Maka dapat disimpulkan diantara variabel independen terdapat masalah multikolinearitas.

Pendeteksian multikolinearitas juga dapat dilakukan dengan matriks korelasi Pearson antarvariabel independen. Apabila nilai determinan matriks korelasi sama dengan nol, maka telah terjadi masalah multikolinearitas. Matriks **K** yang merupakan matriks korelasi berasal dari Tabel 4.2. disajikan sebagai berikut.

$$K = \begin{bmatrix} 1 & 0,407 & 0,661 & 0,570 & 0,743 & 0,322 & 0,545 & 0,747 & -0,310 \\ 0,407 & 1 & -0,034 & 0,155 & -0,053 & -0,111 & 0,524 & 0,805 & 0,805 \\ 0,661 & -0,034 & 1 & 0,570 & 0,792 & 0,581 & 0,624 & 0,447 & 0,082 \\ 0,570 & 0,155 & 0,570 & 1 & 0,733 & 0,256 & 0,272 & 0,505 & -0,292 \\ 0,743 & -0,053 & 0,792 & 0,733 & 1 & 0,459 & 0,423 & 0,521 & -0,278 \\ 0,322 & -0,111 & 0,581 & 0,256 & 0,459 & 1 & 0,327 & 0,207 & -0,066 \\ 0,545 & 0,524 & 0,624 & 0,272 & 0,423 & 0,327 & 1 & 0,757 & 0,444 \\ 0,747 & 0,805 & 0,447 & 0,505 & 0,521 & 0,207 & 0,757 & 1 & 0,068 \\ -0,310 & 0,169 & 0,082 & -0,292 & -0,278 & -0,066 & 0,444 & 0,068 & 1 \end{bmatrix}$$

Determinan dari matriks korelasi **K** adalah 0,00002517035 yang lengkapnya terdapat pada Lampiran 2. Hasil determinan tersebut menunjukkan bahwa nilai determinan mendekati nol sehingga mengindikasikan adanya masalah multikolinearitas.

## 2. Menghitung Nilai VIF (*Variance Inflation Factor*)

Mendeteksi adanya multikolinearitas menggunakan metode VIF dilakukan dengan melihat apabila nilai VIF > 5 maka dapat disimpulkan bahwa terdapat multikolinearitas pada data. Nilai VIF untuk data yang ada disajikan pada Tabel 3 sebagai berikut.

**Tabel 3** Nilai VIF Variabel Independen

Variabel	Nilai VIF	Keterangan
X <sub>1</sub>	35,782584	Terdapat multikolinearitas
X <sub>2</sub>	6,090765	Terdapat multikolinearitas
X <sub>3</sub>	2,824175	Tidak terdapat multikolinearitas
X <sub>4</sub>	25,611309	Terdapat multikolinearitas

X <sub>5</sub>	1,696251	Tidak terdapat multikolinearitas
X <sub>6</sub>	6,564255	Terdapat multikolinearitas
X <sub>7</sub>	57,474698	Terdapat multikolinearitas
X <sub>8</sub>	2,394530	Tidak terdapat multikolinearitas

Berdasarkan Tabel 4.3 diketahui bahwa nilai VIF untuk variabel (X<sub>1</sub>), (X<sub>2</sub>), (X<sub>4</sub>), (X<sub>6</sub>) dan (X<sub>7</sub>) lebih dari 5 sehingga disimpulkan terdapat masalah multikolinearitas pada variabel-variabel tersebut. Dengan kata lain variabel-variabel tersebut telah saling mempengaruhi satu sama lain.

**Penerapan Regulasi LASSO Pada Model Regresi**

**1. Menduga Koefisien Regresi LASSO**

Tahapan-tahapan untuk setiap variabel independen yang masuk ke dalam model regresi LASSO dapat dilihat pada Tabel 4 sebagai berikut.

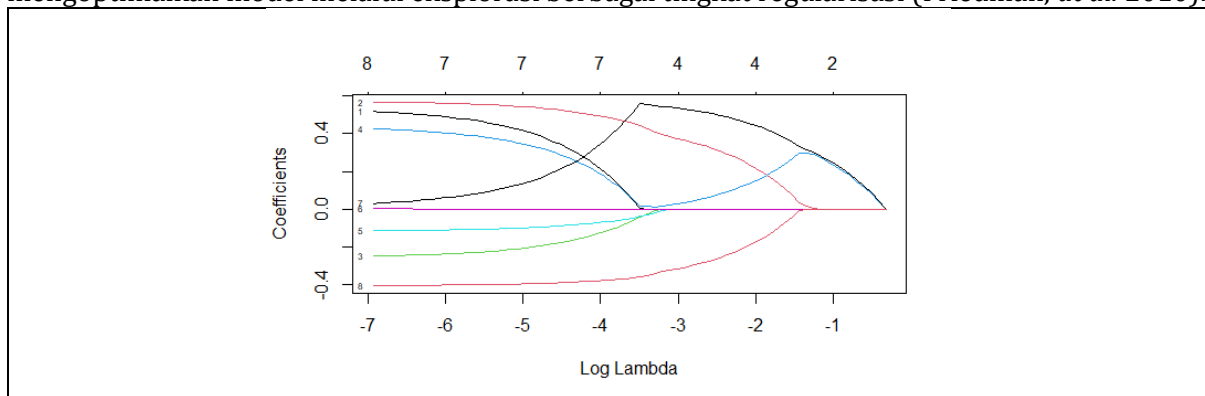
**Tabel 4** Variabel Independen yang Masuk ke Dalam Model pada Setiap Tahapan

Tahapan	Variabel Independen yang Masuk ke Dalam Model										
1	X <sub>7</sub>										
2	X <sub>7</sub>	X <sub>4</sub>									
3	X <sub>7</sub>	X <sub>4</sub>	X <sub>2</sub>								
4	X <sub>7</sub>	X <sub>4</sub>	X <sub>2</sub>	X <sub>8</sub>							
5	X <sub>7</sub>	X <sub>4</sub>	X <sub>2</sub>	X <sub>8</sub>	X <sub>5</sub>						
6	X <sub>7</sub>	X <sub>4</sub>	X <sub>2</sub>	X <sub>8</sub>	X <sub>5</sub>	X <sub>3</sub>					
7	X <sub>7</sub>	X <sub>4</sub>	X <sub>2</sub>	X <sub>8</sub>	X <sub>5</sub>	X <sub>3</sub>	X <sub>1</sub>				
8	X <sub>7</sub>	X <sub>4</sub>	X <sub>2</sub>	X <sub>8</sub>	X <sub>5</sub>	X <sub>3</sub>	X <sub>1</sub>	X <sub>6</sub>			
9	X <sub>7</sub>	X <sub>4</sub>	X <sub>2</sub>	X <sub>8</sub>	X <sub>5</sub>	X <sub>3</sub>	X <sub>1</sub>	X <sub>6</sub>	X <sub>7</sub>		
10	X <sub>7</sub>	X <sub>4</sub>	X <sub>2</sub>	X <sub>8</sub>	X <sub>5</sub>	X <sub>3</sub>	X <sub>1</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>7</sub>	

Berdasarkan Tabel 4.4, diketahui bahwa variabel independen yang masuk ke dalam model secara bertahap ialah X<sub>7</sub>, X<sub>4</sub>, X<sub>2</sub>, X<sub>8</sub>, X<sub>5</sub>, X<sub>3</sub>, X<sub>1</sub>, X<sub>6</sub>, X<sub>7</sub> dan X<sub>7</sub>. Pada tabel di atas, variabel independen yang masuk ke dalam model diurut berdasarkan variabel yang masuk ke dalam model.

**2. Menentukan Lambda Optimum**

Rentang nilai  $\lambda$  memberikan cakupan yang halus di seluruh rentang, bertujuan untuk mengoptimalkan model melalui eksplorasi berbagai tingkat regularisasi (Friedman, *at al.* 2010).



**Gambar 1** Koefisien Terhadap Lambda

Plot penyusutan koefisien yang dihasilkan menunjukkan bagaimana nilai koefisien regresi dari setiap variabel independen berubah seiring dengan meningkatnya nilai lambda dalam regresi LASSO. Pada plot terlihat seiring bertambahnya nilai lambda, sebagian besar koefisien mendekati nol akibat penalti regularisasi yang kuat, menghasilkan model yang lebih sederhana. Sebaliknya, saat lambda mengecil (mendekati nol), penalti berkurang dan lebih banyak koefisien menjadi *non-zero*, mencerminkan model yang lebih kompleks. Garis-garis yang tetap *non-zero* dalam rentang log (lambda) yang luas menunjukkan variabel yang paling berkontribusi dalam model.

Selanjutnya data dibagi menjadi 5-fold, fold pertama dan kedua masing-masing berisi 4 amatan, fold ketiga hingga kelima masing masing berisi 3 amatan.

$$\begin{aligned}
 \text{fold ke-1} = \text{MSE}_1 &= \frac{1}{n_1} \sum_{i=1}^{n_1} (y_{i1} - y_i)^2 \\
 &= \frac{1}{4} (153 - (-0,375))^2 + (778 - 0,868)^2 + (191 - 0,266)^2 + (205 - 0,381)^2 \\
 &\vdots \\
 \text{fold ke-5} = \text{MSE}_5 &= \frac{1}{n_5} \sum_{i=1}^{n_5} (y_{i5} - y_i)^2 \\
 &= \frac{1}{3} (0 - (-0,375))^2 + (1006 - 0,868)^2 + (3 - 0,266)^2
 \end{aligned}$$

Proses akan terus berulang hingga mencapai *fold* terakhir yang digunakan sebagai data validasi. Nilai CV MSE yang dihasilkan dengan mode *step* dapat dilihat pada Tabel 4.5 sebagai berikut.

**Tabel 5** Nilai CV MSE dengan Mode *Step*

Tahap	CV MSE
1	2,990
2	2,954
3	1,199
4	0,999
5	0,180
6	0,058
7	0,069
8	0,072
9	0,067
10	0,061

Berdasarkan Tabel 5 diperoleh nilai CV MSE pada setiap tahapan. Diketahui nilai lambda optimum dapat diambil dari nilai CV MSE dengan rata-rata paling rendah, dalam hal ini diperoleh nilai lambda optimum sebesar 0,058.

### 3. Pemilihan Model Terbaik

Selanjutnya untuk menentukan model terbaik regresi LASSO digunakan validasi silang (CV) dengan menggunakan mode *step*. Pada penelitian ini digunakan validasi silang lipat-5 untuk menghasilkan nilai CV dengan bias yang tinggi tetapi dengan ragam yang rendah. Mode *step* diawali dengan menghitung nilai dari validasi silang untuk setiap tahapan, di mana satu variabel masuk ke dalam model.

Berdasarkan Tabel 5, nilai CV MSE model terbaik dengan menggunakan mode *step* ada pada tahap 6 karena memperoleh nilai CV MSE minimum sebesar 0,058. Adapun variabel independen yang masuk ke dalam model pada tahap 6 ialah  $X_7$ ,  $X_4$ ,  $X_2$ ,  $X_8$ ,  $X_5$ , dan  $X_3$ .

### 4. Evaluasi Performa Model

Evaluasi performa model regresi LASSO dapat dilihat pada tabel berikut.

**Tabel 6** Nilai  $R^2$  Model Regresi LASSO

Metode	$R^2$
LASSO	0,8666

$R^2$  sebesar 86,66% menunjukkan seberapa besar model dalam menjelaskan variabilitas data. Ini berarti bahwa model memiliki kemampuan prediksi yang cukup baik dengan menyisakan 6 variabel independen. Namun, ada sebesar 13,34% variabilitas yang tidak dapat dijelaskan oleh model. Hal ini berarti ada faktor-faktor lain yang mempengaruhi jumlah kejahatan yang dilaporkan yang tidak ditangkap oleh model.

### 5. Interpretasi Model

Diperoleh persamaan model regresi LASSO dengan nilai lambda optimum sebagai berikut.

$$Y = -9,870 \times 10^{-6} + 0,406X_2 - 0,374X_3 + 0,009X_4 - 0,015X_5 + 0,048X_7 - 0,339X_8$$

Interpretasi:

- Konstanta ( $\beta_0$ ) sebesar  $-9,870 \times 10^{-6}$  menunjukkan bahwa ketika semua variabel independen dianggap konstan atau bernilai 0, maka jumlah kejahatan yang dilaporkan (*crime total*) diperkirakan sebesar  $-0,00000987$  atau  $-9,870 \times 10^{-6}$ .

2. Jika variabel rata-rata lama sekolah ( $X_2$ ) bertambah sebanyak dua belas tahun sedangkan tingkat pengangguran terbuka ( $X_3$ ), kepadatan penduduk ( $X_4$ ), gini rasio ( $X_5$ ), jumlah penduduk ( $X_7$ ) dan rasio jenis kelamin ( $X_8$ ) konstan atau tidak mengalami perubahan maka jumlah kejahatan yang dilaporkan ( $Y$ ) akan meningkat sebanyak 5 kasus. Meskipun secara umum hal tersebut sedikit kontradiktif, peningkatan rata-rata lama sekolah justru bisa berkorelasi dengan naiknya jumlah kejahatan yang dilaporkan. Hal ini dapat terjadi karena masyarakat yang lebih terdidik cenderung memiliki kesadaran hukum yang lebih tinggi, sehingga lebih aktif melaporkan tindak kejahatan yang sebelumnya mungkin diabaikan. Selain itu, tingkat pendidikan yang lebih tinggi juga dapat memunculkan jenis kejahatan intelektual seperti penipuan atau kejahatan siber.
3. Jika variabel tingkat pengangguran terbuka ( $X_3$ ) bertambah sebanyak sepuluh satuan sedangkan rata-rata lama sekolah ( $X_2$ ), kepadatan penduduk ( $X_4$ ), gini rasio ( $X_5$ ), jumlah penduduk ( $X_7$ ) dan rasio jenis kelamin ( $X_8$ ) konstan atau tidak mengalami perubahan maka jumlah kejahatan yang dilaporkan ( $Y$ ) akan mengalami penurunan sebanyak 3 kasus.
4. Jika variabel kepadatan penduduk ( $X_4$ ) bertambah sebanyak seribu jiwa sedangkan rata-rata lama sekolah ( $X_2$ ), tingkat pengangguran terbuka ( $X_3$ ), gini rasio ( $X_5$ ), jumlah penduduk ( $X_7$ ) dan rasio jenis kelamin ( $X_8$ ) konstan atau tidak mengalami perubahan maka jumlah kejahatan yang dilaporkan ( $Y$ ) akan meningkat sebanyak 9 kasus.
5. Jika variabel gini rasio ( $X_5$ ) bertambah sebanyak seratus satuan sedangkan rata-rata lama sekolah ( $X_2$ ), tingkat pengangguran terbuka ( $X_3$ ), kepadatan penduduk ( $X_4$ ), jumlah penduduk ( $X_7$ ) dan rasio jenis kelamin ( $X_8$ ) konstan atau tidak mengalami perubahan maka jumlah kejahatan yang dilaporkan ( $Y$ ) akan mengalami penurunan sebanyak 1 kasus. Umumnya, ketimpangan pengeluaran (gini rasio) dapat meningkatkan potensi kejahatan. Namun, hasil ini bisa saja terjadi karena rendahnya pelaporan kejahatan akibat kurangnya kepercayaan masyarakat terhadap aparat penegak hukum.
6. Jika variabel jumlah penduduk ( $X_7$ ) bertambah sebanyak seribu jiwa sedangkan rata-rata lama sekolah ( $X_2$ ), tingkat pengangguran terbuka ( $X_3$ ), kepadatan penduduk ( $X_4$ ), gini rasio ( $X_5$ ) dan rasio jenis kelamin ( $X_8$ ) konstan atau tidak mengalami perubahan maka jumlah kejahatan yang dilaporkan ( $Y$ ) akan meningkat sebanyak 48 kasus.
7. Jika variabel rasio jenis kelamin ( $X_8$ ) naik sepuluh satuan sedangkan rata-rata lama sekolah ( $X_2$ ), tingkat pengangguran terbuka ( $X_3$ ), kepadatan penduduk ( $X_4$ ), gini rasio ( $X_5$ ) dan jumlah penduduk ( $X_7$ ) konstan atau tidak mengalami perubahan maka jumlah kejahatan yang dilaporkan ( $Y$ ) akan mengalami penurunan sebanyak 3 kasus. Secara umum laki-laki cenderung lebih berpotensi melakukan tindak kriminal sehingga secara teoritis peningkatan rasio laki-laki seharusnya berkorelasi positif dengan jumlah kejahatan yang dilaporkan. Namun, hasil negatif ini dapat disebabkan oleh kemungkinan rendahnya tingkat pelaporan kejahatan. Faktor sosial dan budaya seperti norma patriarkis atau keterbatasan akses pelaporan, menyebabkan sebagian besar kejahatan tidak tercatat secara resmi.

Berdasarkan model tersebut, diketahui bahwa variabel independen yang berpengaruh secara nyata terhadap variabel jumlah kejahatan yang dilaporkan (*crime total*) ialah rata-rata lama sekolah ( $X_2$ ), tingkat pengangguran terbuka ( $X_3$ ), kepadatan penduduk ( $X_4$ ), gini rasio ( $X_5$ ), jumlah penduduk ( $X_7$ ) dan rasio jenis kelamin ( $X_8$ ).

Untuk melihat lebih jelas nilai VIF dengan menggunakan LASSO ditampilkan dalam tabel 7 sebagai berikut.

**Tabel 7** Nilai VIF dengan Metode LASSO

Variabel	VIF
$X_2$	4,472177
$X_3$	2,433227
$X_4$	4,682550
$X_5$	1,557731
$X_7$	1,484632
$X_8$	1,626426

Setelah diperoleh model regresi LASSO, selanjutnya dilakukan pengecekan apakah data masih mengandung multikolinearitas. Hal ini dapat dilihat pada tabel 7 di atas, terlihat nilai VIF dari enam variabel independen  $< 5$  maka disimpulkan sudah tidak ada multikolinearitas pada data. Sehingga dengan ini dapat dibuktikan bahwa regresi LASSO dapat digunakan untuk mengatasi permasalahan multikolinearitas.

## D. Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, penelitian ini berhasil menerapkan metode LASSO dalam menangani permasalahan multikolinearitas pada model regresi linear berganda yang digunakan untuk memprediksi Jumlah Kejahatan yang Dilaporkan (*Crime Total*) di Provinsi Sulawesi Tenggara. Analisis menunjukkan bahwa:

1. Diperoleh nilai CV MSE untuk  $\lambda$  optimum sebesar 0,058 yang diterapkan pada penyederhanaan model membuat koefisien menyusut lebih stabil sehingga telah mengatasi multikolinearitas. Dengan nilai  $R^2$  sebesar 86,66% menunjukkan model yang dihasilkan dapat digunakan cukup akurat dalam memprediksi Jumlah Kejahatan yang Dilaporkan (*Crime Total*) di Provinsi Sulawesi Tenggara.
2. Model LASSO yang terbentuk adalah sebagai berikut:

$$Y = -9,870 \times 10^{-6} + 0,406X_2 - 0,374X_3 + 0,009X_4 - 0,015X_5 + 0,048X_7 - 0,339X_8$$

## Saran

Berdasarkan hasil penelitian dan pembahasan, dibuktikan bahwa LASSO merupakan metode yang baik dalam mengatasi masalah multikolinearitas. Oleh karena itu, penulis menyarankan menggunakan metode yang lain untuk dijadikan pembandingan seperti Regresi *Ridge*, *Elastic Net*, dan lain sebagainya.

## E. Referensi

- Arum, P. R., Haris, M. A. 2019. Analisis Faktor-Faktor yang Mempengaruhi Jumlah Penduduk di Kota Semarang Menggunakan Metode Regresi Data Panel. *Jurnal Statistika*, Vol 12(2): 36-41.
- BPS. 2022. *Rata-rata Upah/Gaji Bersih Sebulan Buruh/Karyawan/Pegawai Menurut Provinsi dan Jenis Pekerjaan Utama, 2022*. Jakarta: Badan Pusat Statistik.
- BPS. 2022. *Statistik Kejahatan yang Dilaporkan (Crime Total) Provinsi Sulawesi Tenggara 2022*. Kendari: Badan Pusat Statistik Provinsi Sulawesi Tenggara.
- BPS. 2024. *Rata-rata Lama Sekolah (RLS)*. Karanganyar: Badan Pusat Statistik Kabupaten Karanganyar.
- Desinta. 2022. Faktor-faktor yang Memengaruhi Kejadian Kejahatan di Provinsi Jawa Barat Tahun 2018-2020. *Jurnal Ilmiah Populer*, 5.
- Efron, B., Tibshirani, R. 1993. *Book an Intoduction to The Bootstrap*. Chapman and Hall/CRC: New York.
- Fatihah, A.I. 2022. *Penerapan Metode Jackknife Ridge Regression dalam Penanganan Multikolinearitas (Studi Kasus: Tingkat Kemiskinan di Kabupaten Grobogan Tahun 2011-2021)*. Universitas Islam Negeri Walisongo
- Firdaus & Zamal. 2009. *Korelasi antara Pelatihan Teknis Perpajakan, Pengalaman dan Motivasi Pemeriksa Pajak dengan Kinerja Pemeriksa Pajak pada Kantor Pelayanan Pajak di Jakarta Barat*. Fakultas Ekonomi dan Ilmu Sosial Universitas Islam Negeri Syarif Hidayatullah Jakarta
- Hastie, T., Tibshirani, R., & Friedman, J. 2015. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.).
- James, G., Witten, D., Hastie, T., & Tibshirani, R. 2013. *An Introduction to Statistical Learning with Applications in R* (G. Casella, S. Fienberg, & I. Lestari, P.S., Martha, S., & Debatara, N.N. 2022. Penerapan Metode Regresi Ridge Pada Kasus Angka Kematian Bayi di Provinsi Jawa Timur. *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, Vol 11(4), 603-610.
- Ningsih, D.R., Intan, P.K., & Yuliati, D. 2023. Pemodelan Tindak Pidana Kriminalitas di Kota Tangerang Menggunakan Metode Regresi LASSO. *Journal of Statistics and Its Application*, Vol 4(1).
- Nurdin, I. Sugiman, & Sunarmi. 2018. Penerapan Kombinasi Metode Ridge Regression (RR) dan Metode Generalized Least Square (GLS) untuk Mengatasi Masalah Multikolinearitas dan Autokorelasi. *Jurnal MIPA*, Vol 41(2): 58-68.
- Pendi. 2021. Analisis Regresi dengan Metode Komponen Utama dalam Mengatasi Masalah Multikolinearitas. *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, Vol 10(1): 131-138.
- Priatna. 2015. *Analisis Pengaruh Sosial Ekonomi Terhadap Tingkat Kriminalitas Pencurian di Daerah Istimewa Yogyakarta Tahun 2010-2015*. Skripsi Fakultas Ekonomi, Universitas Muhammadiyah Yogyakarta: Yogyakarta.
- Robbani, M., Agustiani, F., & Herhhyanto, N. 2019. Regresi *Least Absolute Shrinkage and Selection Operator* (Lasso) Pada Kasus Inflasi di Indonesia Tahun 2014-2017. *Jurnal Euramatika*, Vol 7(2).

- Sulantari, S., Hariadi, W., Putra, E.D., & Anas, A. 2022. Analisis Regresi Linier Berganda Untuk Memodelkan Faktor Yang Mempengaruhi Nilai Penambahan Utang Tahunan Negara Indonesia. *Unisda Journal of Mathematics and Computer*, 10(1), 36-46.
- Supriyadi, E., Scolatika, M., & Sugiman. 2017. Perbandingan Metode *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk Mengatasi Masalah Multikolinearitas pada Model Regresi Linear Berganda. *UNNES Journal of Mathematics*, Vol 6(2).
- Wasailane, T.L., Talakua, M.W., & Lesnussa, Y.A. 2014. Model Regresi Ridge untuk Mengatasi Model Regresi Linear Berganda yang Mengandung Multikolinearitas (Studi Kasus: Data Pertumbuhan Bayi di Kelurahan Namaelo RT 001, Kota Masohi). *Jurnal Barekeng*, Vol 8(1): 31-37.