

## Perbandingan Metode Regresi *Ridge* dan *Jackknife Ridge Regression* pada Data Tingkat Pengangguran Terbuka

Agita Andini<sup>1</sup>, Etis Sunandi<sup>2\*</sup>, Pepi Novianti<sup>3</sup>, Idhia Sriliana<sup>4</sup>, Winalia Agwil<sup>5</sup>  
<sup>1,2,3,4,5</sup>Program Studi Statistika, Universitas Bengkulu  
e-mail: [esunandi@unib.ac.id](mailto:esunandi@unib.ac.id)

*Diajukan: 2 Oktober 2024, Diperbaiki: 30 Oktober 2024, Diterima: 20 September 2024*

### Abstrak

Analisis regresi merupakan teknik statistik yang digunakan untuk menganalisis hubungan antara variabel prediktor dan variabel respon. Salah satu metode estimasi parameter yang sering digunakan untuk analisis regresi adalah *Ordinary Least Square*. Metode ini menghasilkan pendugaan yang tidak bias dan efisien, yang dikenal sebagai BLUE (*Best Linear Unbiased Estimator*). Pada analisis regresi linear berganda yang melibatkan lebih dari satu variabel prediktor, penting untuk memenuhi asumsi model seperti tidak adanya multikolinearitas. Multikolinearitas merupakan kondisi di mana variabel prediktor memiliki korelasi yang tinggi satu sama lain, sehingga dapat mengganggu stabilitas estimasi parameter. Oleh karena itu, untuk mengatasi masalah tersebut digunakan metode Regresi *Ridge* dan *Jackknife Ridge Regression*. Kedua metode ini merupakan hasil modifikasi metode kuadrat terkecil dengan cara penambahan nilai konstanta bias. Penelitian ini menggunakan data Tingkat Pengangguran Terbuka (TPT) di Pulau Sumatra tahun 2022, dan terdapat 3 variabel prediktor yang memiliki multikolinearitas. Berdasarkan analisis yang dilakukan dengan membandingkan nilai *Mean Squared Error* (MSE), metode *Jackknife Ridge Regression* menghasilkan nilai MSE terkecil, yaitu sebesar 0,003. Kedua metode tersebut efektif mengatasi multikolinearitas dan mengidentifikasi variabel prediktor signifikan terhadap TPT di Pulau Sumatra, yaitu Indeks Pembangunan Manusia (IPM), rata-rata lama sekolah, jumlah penduduk miskin, Umur Harapan Hidup (UHH), kepadatan penduduk dan penduduk nonaktif.

**Kata Kunci:** Analisis Regresi, Multikolinearitas, Pengangguran.

### Abstract

*Regression analysis is a statistical technique used to analyze the relationship between predictor and response variables. One of the parameter estimation methods commonly used for regression analysis is Ordinary Least Squares. This method produces unbiased and efficient estimates, known as BLUE (Best Linear Unbiased Estimator). In multiple linear regression analysis involving more than one predictor variable, it is essential to meet model assumptions such as the absence of multicollinearity. Multicollinearity is a condition where predictor variables have a high correlation, which can disrupt the stability of parameter estimates. Therefore, Ridge Regression and Jackknife Ridge Regression methods were used to address this issue. Both methods modify the least squares method by adding a bias constant value. This research uses the Open Unemployment Rate (OUR) data in Sumatra in 2022, and 3 predictor variables exhibit multicollinearity. Based on the analysis comparing the Mean Squared Error (MSE) values, the Jackknife Ridge Regression method yields the smallest MSE value, 0.004. Both methods are effective in addressing multicollinearity and identifying significant predictor variables for OUR in Sumatra Island, namely the Human Development Index (HDI), average years of schooling, number of poor people, Life Expectancy (LE), population density and inactive population.*

**Keywords:** Regression Analysis, Multicollinearity, Unemployment

## 1 Pendahuluan

Analisis regresi merupakan teknik statistik yang digunakan untuk memahami hubungan antar variabel [1]. Regresi linear terbagi menjadi dua jenis yaitu sederhana (dengan satu prediktor) dan berganda (dengan dua atau lebih prediktor). Model regresi yang baik menghasilkan kesalahan estimasi yang minimum. Metode kuadrat terkecil atau *Ordinary Least Square* sering digunakan karena memberikan estimasi linear terbaik yang tidak bias dan bersifat *Best Linear Unbiased Estimator* (BLUE) [1], [2].

Salah satu asumsi penting dalam regresi OLS adalah tidak adanya multikolinearitas. Multikolinearitas terjadi ketika terdapat hubungan linear antara variabel prediktor dalam regresi berganda [3]. Hal ini dapat mengakibatkan model kesulitan menjelaskan hubungan antara prediktor dan respons secara akurat. Selain itu, multikolinearitas meningkatkan varian parameter estimasi, yang menurunkan akurasi estimasi [4].

Salah satu metode yang dapat digunakan untuk mengatasi multikolinearitas adalah Regresi *Ridge*. Regresi *Ridge* awalnya diperkenalkan oleh Hoerl dan R.W Kennard pada tahun 1962 sebagai solusi untuk mengatasi masalah multikolinearitas dengan mengenalkan nilai tetapan bias  $c$  yang ditambahkan ke matriks  $X^tX$ . Metode Regresi *Ridge* diperoleh dengan cara dengan meminimumkan jumlah kuadrat sisa dan dilakukan penambahan konstanta bias agar nilai koefisien berkurang dan mendekati nol [3].

Kelemahan dari metode Regresi *Ridge* yaitu nilai bias yang dihasilkan tidak selalu kecil. Oleh karena itu, *Jackknife Ridge Regression* (JRR) dikembangkan dengan tujuan untuk mengurangi nilai bias dari hasil estimasi Regresi *Ridge* [5]. Metode JRR dapat diperoleh dengan menggabungkan metode *Generalized Ridge Regression* dan resampling *jackknife* [6]. Metode *Generalized Ridge Regression* merupakan metode hasil dari pengembangan Regresi *Ridge* dengan cara modifikasi penambahan konstanta bias  $c$  yang berbeda untuk masing-masing variabel prediktornya [3]. Pendekatan pada metode JRR memberikan bobot yang berbeda untuk setiap observasi secara berulang, sehingga memungkinkan model untuk lebih responsif terhadap fluktuasi dalam data [5].

Permasalahan multikolinearitas juga dapat muncul dalam analisis regresi yang menggunakan data-data ekonomi, seperti data tingkat pengangguran terbuka. Hal ini disebabkan karena data-data ekonomi sering berkorelasi. Dalam penelitian [7] metode *Generalized Ridge Regression* digunakan untuk mengatasi multikolinearitas pada data tingkat pengangguran terbuka di Indonesia. Selain itu, pada penelitian [8] menerapkan metode *Elastic Net* untuk mengatasi permasalahan multikolinearitas pada data tingkat pengangguran terbuka di Jawa Barat.

Pada tahun 2022, tingkat pengangguran di berbagai provinsi di Pulau Sumatra bervariasi. Sebagai contoh di Sumatera Barat, Tingkat Pengangguran Terbuka (TPT) pada Agustus 2022 tercatat sebesar 6,28%, sedikit menurun dibandingkan tahun sebelumnya [9]. Secara keseluruhan, meskipun ada sedikit penurunan di beberapa provinsi, tingkat pengangguran masih cukup tinggi di beberapa wilayah di Sumatra pada tahun 2022. Oleh karena itu, Penelitian ini melakukan perbandingan antara metode Regresi *Ridge* dan *Jackknife Ridge Regression* untuk menentukan metode yang lebih efektif dalam mengatasi multikolinearitas, terutama pada data ekonomi dengan jumlah pengamatan yang lebih banyak dibandingkan dengan penelitian sebelumnya [7]. Sehingga penelitian ini diharapkan dapat mengidentifikasi metode yang menghasilkan estimasi lebih akurat dan sesuai untuk diterapkan pada data ekonomi yang kompleks.

## 2 Metode Penelitian

Penelitian ini menggunakan data sekunder yang diperoleh melalui *website* Badan Pusat Statistik (BPS) dan publikasi Provinsi dalam Angka tahun 2023 [9]. Jumlah pengamatan pada data yang digunakan sebanyak 154 observasi berdasarkan kabupaten/kota di Pulau Sumatra. Keterangan variabel pada penelitian ini disajikan pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Kode	Keterangan	Skala dan Satuan
Variabel respon			
	y	Tingkat Pengangguran Terbuka	Rasio (%)
Variabel prediktor			
	$x_1$	Indeks Pembangunan Manusia	Rasio (indeks)
	$x_2$	Rata-Rata Lama Sekolah	Rasio (tahun)
	$x_3$	Jumlah Penduduk Miskin	Diskrit (ribu jiwa)
	$x_4$	Umur Harapan Hidup	Rasio (tahun)
	$x_5$	Laju PDRB	Rasio (%)
	$x_6$	Jumlah Penduduk	Diskrit (jiwa)
	$x_7$	Pengeluaran Perkapita	Rasio (ribu rupiah)
	$x_8$	Kepadatan Penduduk	Rasio (jiwa/km <sup>2</sup> )
	$x_9$	Penduduk Nonaktif	Diskrit (jiwa)

### 2.1 Regresi *Ridge*

Regresi *Ridge* adalah metode estimasi parameter yang cenderung mengurangi bias dan varian, menghasilkan parameter yang lebih stabil dan lebih akurat daripada metode *Ordinary Least Square* yang cenderung tidak mengurangi bias [3]. Penduga parameter regresi *Ridge* yaitu :

$$\hat{\beta}_R = (X^{*t}X^* + cI)^{-1}X^{*t}Y^* \quad (1)$$

Di mana  $X^*_{(n \times k)}$  merupakan matriks variabel prediktor setelah distandarisasi,  $Y^*$  adalah vektor variabel respon setelah distandarisasi,  $c$  adalah nilai konstanta bias dan  $I_{(k \times k)}$  dinyatakan sebagai matriks identitas.

2.2 Jackknife Ridge Regression

Metode *Jackknife Ridge Regression* dikembangkan dari *Generalized Ridge Regression* dengan menerapkan teknik resampling *jackknife* pada metode tersebut. Oleh karena itu, metode ini dikenal sebagai *Jackknife Ridge Regression*. Hasil dari dilakukannya modifikasi tersebut, memperoleh persamaan sebagai berikut [5]:

$$Y_{(-i)} = Z_{(-i)}\hat{\gamma}_{GRR} + \epsilon^* \tag{2}$$

Di mana  $Y_{(-i)}$  merupakan vektor variabel respon dengan pengamatan ke- $i$  dihapus,  $Z_{(-i)}$  merupakan matriks  $Z_{(n \times k)}$  dengan setiap baris ke- $i$  dihapus. Matriks  $Z_{(n \times k)}$  diperoleh dari hasil perkalian matriks  $X_{(n \times k)}$  dan matriks  $C$ . Matriks  $C_{(k \times k)}$  adalah matriks orthogonal dari vektor eigen matriks  $X^t X$  dan  $\epsilon^*$  adalah vektor residual dengan pengamatan ke- $i$  dihapus.

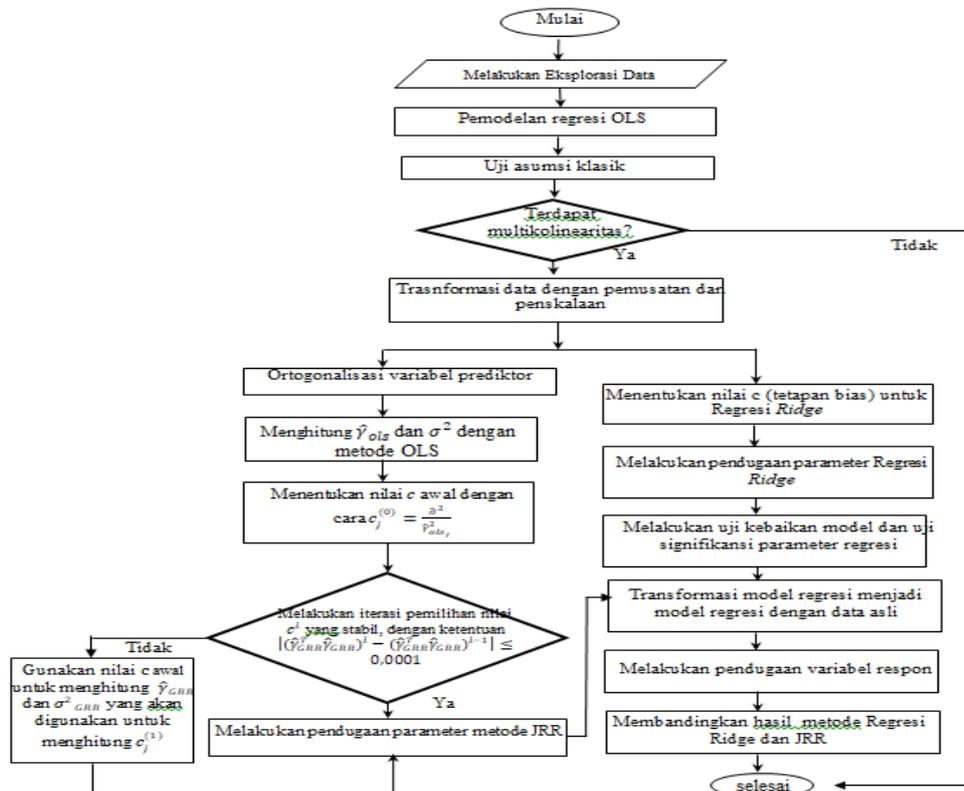
Parameter  $\hat{\beta}$  untuk *Jackknife Ridge Regression* dapat diperoleh dengan menggunakan persamaan berikut :

$$\hat{\beta}_{JRR} = C\hat{\gamma}_{JRR} \tag{3}$$

Di mana  $\hat{\beta}_{JRR}$  adalah vektor kolom estimasi koefisien regresi metode JRR dan  $\hat{\gamma}_{JRR}$  adalah vektor kolom estimasi parameter  $\gamma$  dengan regresi JRR.

2.3 Analisis Data

Langkah-langkah yang dilakukan pada penelitian ini adalah sebagai berikut :



Gambar 1. Flowchart Tahapan Analisis

### 3 Hasil dan Pembahasan

#### 3.1 Analisis Regresi Metode *Ordinary Least Square*

Pada tahapan pertama dilakukan estimasi parameter menggunakan Regresi OLS dengan menggunakan data awal. Pada Tabel 2 disajikan hasil estimasi parameter Regresi OLS. Diperoleh model regresi linear berganda dengan seluruh variabel prediktor seperti berikut :

$$\hat{y}_{ols} = 19,005 + 2,292x_1 - 0,056x_2 + 0,005x_3 - 0,484x_4 + 0,083x_5 + 9,57 \times 10^{-7}x_6 - 3,20 \times 10^{-5}x_7 + 1,69 \times 10^{-4}x_8 + 3,77 \times 10^{-7}x_9 \quad (4)$$

#### 3.2 Uji Asumsi Klasik

Pada Tabel 4 disajikan hasil uji normalitas dengan menggunakan uji *Jarque Bera*. Diperoleh hasil bahwa dengan menggunakan  $\alpha = 5\%$ , diperoleh  $P_{value} = 0,586 > 0,05$ . Sehingga keputusannya terima  $H_0$  yang artinya residual berdistribusi normal.

Tabel 2. Hasil Uji Normalitas

<i>JB</i>	<i>pvalue</i>	Keterangan
1,070	0,586	Terima $H_0$

Selanjutnya, dilakukan uji heteroskedastisitas dengan menggunakan uji *Breusch Pagan*. Pada Tabel 3 diperoleh bahwa dengan menggunakan  $\alpha = 5\%$  diperoleh  $P_{value} = 0,187 > 0,05$ . Sehingga terima  $H_0$  yang artinya tidak terdapat heteroskedastisitas.

Tabel 3. Hasil Uji Heteroskedastisitas

$\phi$	<i>pvalue</i>	Keterangan
12,489	0,187	Terima $H_0$

#### 3.3 Identifikasi Multikolinearitas

Multikolinearitas dapat diidentifikasi menggunakan nilai VIF dan *tolerance*. Pada Tabel 4 terlihat bahwa terdapat beberapa variabel prediktor yang memiliki nilai  $VIF > 10$  dan nilai  $tolerance < 0,1$  yang artinya pada variabel tersebut mengandung masalah multikolinearitas [10], [11].

Tabel 4. Hasil Identifikasi Multikolinearitas

Variabel	VIF	<i>Tolerance</i>	Keterangan
$x_1$	14,83	0,07	Terjadi Multikolinearitas
$x_2$	8,22	0,12	Tidak Terjadi Multikolinearitas
$x_3$	4,55	0,22	Tidak Terjadi Multikolinearitas
$x_4$	1,80	0,56	Tidak Terjadi Multikolinearitas
$x_5$	1,11	0,90	Tidak Terjadi Multikolinearitas
$x_6$	36,83	0,03	Terjadi Multikolinearitas
$x_7$	3,35	0,30	Tidak Terjadi Multikolinearitas
$x_8$	2,13	0,47	Tidak Terjadi Multikolinearitas
$x_9$	37,85	0,03	Terjadi Multikolinearitas

3.4 Estimasi Parameter Regresi *Ridge*

Pada penelitian ini terdapat dua jenis metode yang digunakan dalam menentukan  $c$  yaitu dua metode yaitu metode (Hoerl, Kennard & Baldwin, 1975) dan metode (Lawless & Wang, 1976) [12].

Tabel 5. Nilai Konstanta Bias Regresi *Ridge*

Metode	Nilai Konstanta Bias
Hoerl, Kennard & Baldwin (1975)	0,04292
Lawless & Wang (1976)	0,06780

Selanjutnya, dilakukan estimasi parameter sehingga diperoleh model regresi *Ridge* dengan data transformasi sebagai berikut :

$$y_{R(hkb)}^* = 0,3251x_1^* + 0,0959x_2^* + 0,0293x_3^* - 0,3084x_4^* + 0,0071x_5^* + 0,0666x_6^* + 0,0206x_7^* + 0,3164x_8^* + 0,2521x_9^* \quad (5)$$

$$y_{R(lw)}^* = 0,2934x_1^* + 0,1162x_2^* + 0,0334x_3^* - 0,2876x_4^* + 0,0093x_5^* + 0,0874x_6^* + 0,0274x_7^* + 0,3055x_8^* + 0,2235x_9^* \quad (6)$$

3.5 Estimasi Parameter *Jackknife Ridge Regression*

Langkah awal yang harus dilakukan adalah melakukan ortogonalisasi variabel-variabel prediktor yang sudah distandarisisasi [13] dan setelah itu menentukan nilai konstanta bias yang disajikan pada Tabel 6.

Tabel 6. Nilai Konstanta Bias Penduga JRR

$c_j^{(1)}$	$\hat{Y}_{GRR}^{(1)}$
0,042	-0,302
7,477	0,019
791,326	$4,959 \times 10^{-5}$
1,377	0,047
0,022	0,426
0,137	0,190
0,088	-0,237
0,107	0,179
0,072	0,130

Berdasarkan hasil estimasi parameter tersebut diperoleh model regresi pada metode *Jackknife Ridge Regression* dengan data standarisasi adalah sebagai berikut :

$$y_{JRR}^* = 0,341x_1^* + 0,101x_2^* + 0,006x_3^* - 0,337x_4^* - 0,010x_5^* + 0,051x_6^* + 0,058x_7^* + 0,309x_8^* + 0,273x_9^* \quad (7)$$

3.6 Perbandingan Hasil Regresi *Ridge* dan *Jackknife Ridge Regression*

Pada Tabel 7 disajikan perbandingan uji signifikansi parameter secara parsial yang disajikan pada tabel berikut. Diperoleh hasil bahwa terdapat perbedaan hasil uji signifikansi parameter setelah dilakukan penerapan metode regresi *Ridge* dan JRR.

Tabel 7. Perbandingan Hasil Uji Parsial

Variabel	Metode JRR	Metode Regresi Ridge HKB 1975	Metode Regresi Ridge LW 1976	Regresi OLS
$x_1$	10,539*	2,784*	3,066*	3,463*
$x_2$	5,916*	0,969	1,370	1,496
$x_3$	38,294*	0,297	0,380	0,723
$x_4$	-8,315*	-4,256*	-4,184*	-2,916*
$x_5$	-0,134	0,117	0,159	0,989
$x_6$	0,525	0,675	1,200	1,382
$x_7$	0,438	0,283	0,398	0,639
$x_8$	2,110*	3,951*	4,042*	1,378
$x_9$	1,872	2,540*	2,925*	1,211

Keterangan : \* = signifikan pada taraf 5%

Pada Tabel 8 diperoleh hasil bahwa metode *Jackknife Ridge Regression* menghasilkan nilai *MSE* yang paling kecil yaitu sebesar 0,003 dan menghasilkan nilai  $R^2$  dan *adjusted R<sup>2</sup>* yang paling besar di mana nilai  $R^2$  yang diperoleh adalah sebesar 0,465. Artinya variabel prediktor mampu menjelaskan keragaman variansi pada variabel respon sebesar 46,5% dan sisanya dipengaruhi oleh faktor lain.

Tabel 8. Perbandingan nilai *MSE*,  $R^2$  dan *adjusted R<sup>2</sup>*

Metode	<i>MSE</i>	$R^2$	<i>Adjusted R<sup>2</sup></i>
Hoerl, Kennard & Baldwin (1975)	0,256	0,441	0,410
Lawless & Wang (1976)	0,280	0,426	0,394
JRR	0,003	0,465	0,432

#### 4 Simpulan

Berdasarkan analisis yang sudah dilakukan, pada metode Regresi *Ridge* faktor-faktor yang mempengaruhi tingkat pengangguran terbuka di Pulau Sumatra adalah variabel IPM, UHH, kepadatan penduduk dan penduduk nonaktif. Sedangkan pada metode *Jackknife Ridge Regression* menghasilkan variabel IPM, RLS, jumlah penduduk miskin, UHH dan kepadatan penduduk. Metode *Jackknife Ridge Regression* menghasilkan nilai *MSE* terkecil yaitu sebesar 0,003 serta nilai  $R^2$  dan *adjusted R<sup>2</sup>* tertinggi.

Berdasarkan hasil penelitian ini, pemerintah diharapkan dapat meningkatkan kualitas pendidikan serta pemerataan kepadatan penduduk agar tingkat pengangguran terbuka di Pulau Sumatra dapat lebih rendah. Kemudian, untuk penelitian selanjutnya dapat menggunakan metode yang lain seperti metode Kibria dan Muniz untuk menentukan nilai konstanta bias. Selain itu, pemilihan metode untuk mengatasi masalah multikolinearitas dapat disesuaikan dengan jumlah data dan tujuan penelitian.

## 5 Daftar Pustaka

- [1] M. H. Kutner, C. J. Nachtsheim, J. Neter, W. Li. *Applied Linear Statistical Models*. Vol. 29. Fifth Edit. New York: Mc Graw Hill, 2005.
- [2] D. C. Montgomery, E. A. Peck, G. Vining. *Introduction to Linear Regression Analysis*. Vol. 11. Fift Editi. New York, 2012.
- [3] T. Hastie, R. Tibshirani, J. Friedman. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. Vol. 26. Second Edi. New York, 2009.
- [4] S. Noora. "Detecting Multicollinearity in Regression Analysis." *American Journal of Applied Mathematics and Statistics* 8(2):39–42. doi: 10.12691/ajams-8-2-1. 2020.
- [5] Tinungki, G. M. "Orthogonal Iteration Process of Determining K Value on Estimator of Jackknife Ridge Regression Parameter." *Journal of Physics: Conference Series* 1341(9). doi: 10.1088/1742-6596/1341/9/092001. 2019.
- [6] F. Adji. "Comparison of Parameter Estimator Efficiency Levels of Path Analysis with Bootstrap and Jack Knife (Delete-5) Resampling Methods on Simulation Data." *Jurnal Matematika, Statistika Dan Komputasi* 16(3):353. doi: 10.20956/jmsk.v16i3.9723. 2020.
- [7] A. Nita, K. Dadan, and N. N. Debararaja. "Metode Generalized Ridge Regression Dalam Mengatasi Multikolinearitas." *Bimaster : Buletin Ilmiah Matematika, Statistika Dan Terapannya* 8(4):679–86. doi: 10.26418/bbimst.v8i4.35879. 2019.
- [8] H. Astri, and L. Wachidah. "Metode Regresi Elastic-Net Untuk Mengatasi Masalah Multikolinearitas Pada Kasus Tingkat Pengangguran Terbuka Di Provinsi Jawa Barat." *Bandung Conference Series: Statistics* 2(2):459–65. doi: 10.29313/bcss.v2i2.4722. 2022.
- [9] Statistik, Badan Pusat. 2023. "Keadaan Angkatan Kerja Provinsi." 260.
- [10] Destiyani, Eka, Rita Rahmawati, and Suparti Suparti. "Pemodelan Regresi Ridge Robust-Mm Dalam Penanganan Multikolinieritas Dan Pencilan (Studi Kasus : Faktor-Faktor Yang Mempengaruhi AKB Di Jawa Tengah Tahun 2017)." *Jurnal Gaussian* 8(1):24–34. 2019.
- [11] Damodar N. Gujarati, Dawn C. Porte. *Basic Econometrics*. Fifth Edition. 2013.
- [12] Ali, Rahmatia G., and Jaka Nugraha. "Penerapan Metode Regresi Ridge Dalam Mengatasi Masalah Multikolinearitas Pada Kasus Indeks Pembangunan Manusia Di Indonesia Tahun 2017." *Prosiding Sendika* 5(2):239–48. 2019.
- [13] A. A. Huda, D. Ispriyanti, and A. Hoyyi. "Metode Modified Jackknife Ridge Regression Dalam Penanganan Multikolinieritas (Studi Kasus Indeks Pembangunan Manusia Di Jawa Tengah)." *Jurnal Gaussian* 10(1):104–13. doi: 10.14710/j.gauss.v10i1.29922. 2021.