

## PEMODELAN PERSENTASE KEMISKINAN DI INDONESIA DENGAN MENGGUNAKAN METODE REGRESI RIDGE

Kayla Nur Khosyatillah Sudza<sup>1)</sup>, Edrick Setiawan<sup>2)</sup>, Lovina Anabelle Citra<sup>3)</sup>, Margaretha Ohhyver<sup>4)\*</sup>,

<sup>1,2,3)</sup>Jurusan Teknik Informatika dan Statistika, School of Computer Science, Universitas Bina Nusantara, Jakarta Barat, DKI Jakarta, Indonesia

<sup>4)</sup>Jurusan Statistika, School of Computer Science, Universitas Bina Nusantara, Jakarta Barat, DKI Jakarta, Indonesia

\*Penulis korespondensi: kayla.sudza002@binus.ac.id

### ABSTRAK

Indonesia dikenal sebagai negara yang kaya akan keanekaragaman budaya dan sumber daya alam. Akan tetapi, terlepas dari keunggulan tersebut, Indonesia masih berjuang untuk melawan tingkat kemiskinan yang cukup tinggi di berbagai daerah, hal ini dapat terlihat dari persentase kemiskinan di setiap provinsi. Penelitian ini bertujuan untuk mencari model regresi yang dapat menemukan faktor-faktor yang berpengaruh terhadap tingkat kemiskinan di Indonesia. Model regresi yang digunakan pada penelitian ini berupa model regresi ridge, dikarenakan terdapat korelasi yang tinggi antara variabel bebas yang digunakan. Berdasarkan analisis lebih lanjut, didapatkan bahwa adanya pengaruh yang signifikan dari Tingkat Pengangguran Terbuka (TPT), Produk Domestik Regional Bruto (PDRB), Lulus SMA, Lama Sekolah, Rokok, dan Seluler terhadap Persentase Kemiskinan. Selain itu, berdasarkan pengukuran nilai *Akaike Information Criterion* (AIC) dan *Bayesian Information Criterion* (BIC), model regresi ridge memiliki performa yang lebih baik. Dengan menggunakan metode regresi ridge, diharapkan penelitian ini dapat menentukan faktor yang berpengaruh akan tingkat kemiskinan dan hasil penelitian ini dapat digunakan untuk mengatasi tingkat kemiskinan di Indonesia.

Kata-kata kunci: Kemiskinan, Multikolinearitas, Regresi Linear, Regresi Ridge

### ABSTRACT

*Indonesia is known as a country rich in cultural diversity and natural resources. However, despite these advantages, Indonesia is still struggling to fight quite high levels of poverty in various regions, this can be seen from the percentage of poverty in each province. This research aims to find a regression model that can find factors that influence the level of poverty in Indonesia. The regression model used in this research is a ridge regression model, because there is a high correlation between the independent variables used. Based on further analysis, it was found that there is a significant influence from the Open Unemployment Rate (TPT), Gross Regional Domestic Product (GRDP), High School Graduation, Years of Schooling, Cigarettes, and Cell Phones on Poverty Percentage. In addition, based on the measurement of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values, the ridge regression model has better performance. By using the ridge regression method, it is hoped that this research can determine the factors that influence poverty levels and the results of this research can be used to overcome poverty levels in Indonesia.*

*Keywords: Linear Regression, Multicollinearity, Poverty, Ridge Regression*

## Pendahuluan

Indonesia merupakan negara yang terkenal dengan kekayaan keanekaragaman budaya dan sumber daya alamnya. Namun terlepas dari keunggulan tersebut, Indonesia masih berjuang melawan tingkat kemiskinan yang tinggi. Banyak penelitian telah dilakukan untuk mengidentifikasi faktor-faktor yang berkontribusi terhadap tingkat kemiskinan di Indonesia. Berbagai faktor, antara lain jumlah penduduk, PDB riil per kapita, angka melek huruf, dan angka harapan hidup, dipelajari dalam penelitian ini. Penelitian ini telah menunjukkan bahwa angka harapan hidup dan PDB riil per kapita mempunyai dampak yang besar terhadap tingkat kemiskinan (Prasada, Yulhar, dan Rosa, 2020.). Faktor penting lainnya yang diketahui mempengaruhi tingkat kemiskinan di Indonesia adalah tingkat suku bunga. Berdasarkan penelitian yang dilakukan antara tahun 2005 sampai 2015, alasan angka kemiskinan tidak berubah akibat perubahan rasio kredit swasta adalah karena Indonesia memiliki tingkat suku bunga yang lebih tinggi dibandingkan Malaysia (2,98% vs. 7,21%). Selanjutnya, faktor sosial ekonomi menjadi faktor utama dalam menentukan kemiskinan di Indonesia. Penting untuk diingat bahwa kemiskinan tidak tersebar merata di seluruh Indonesia.

Ketidakmerataan ini dapat dilihat dari data pemerintah menunjukkan bahwa Pulau Jawa mengalami tingkat pertumbuhan ekonomi tertinggi, yaitu 59%. Sumatera, Kalimantan, Sulawesi, Bali dan Nusa Tenggara, serta Maluku dan Papua masing-masing menyusul sebesar 21,32%, 8,05%, 6,33%, dan 3,06%. Selain itu, angka kemiskinan di tiap provinsi berbeda-beda secara signifikan. Kepulauan Bangka Belitung merupakan provinsi dengan tingkat kemiskinan terendah sebesar 4,53%, sedangkan Papua memiliki angka kemiskinan tertinggi yaitu sebesar 26,68%. Data ini menunjukkan perlunya dilakukan suatu hal untuk memerangi kemiskinan di lokasi-lokasi tertentu, khususnya di daerah perkotaan dimana angka kemiskinan tinggi. Salah satu hal yang berkontribusi terhadap kemiskinan di perkotaan adalah adanya permukiman kumuh dan padat penduduk (Bakhri dan Fauzi, 2021). Untuk mencegah hal ini, penting bagi pihak berwenang untuk mengatasi akar penyebab kemiskinan, seperti terbatasnya akses terhadap pelayanan kesehatan dan pendidikan.

Terdapat banyak faktor lain yang mempengaruhi Tingkat kemiskinan di Indonesia. Berdasarkan studi yang dilakukan oleh Tisniwati, tingkat melek huruf merupakan faktor yang mempengaruhi tingkat kemiskinan di Indonesia (Prasada, Yulhar, dan Rosa, 2020). Selain itu, terdapat beberapa faktor lain yang berpengaruh terhadap tingkat kemiskinan di Indonesia, antara lain Indeks Pembangunan Manusia dan Tingkat Pengangguran Terbuka (Ristika dan Primandhana, 2021), Produk Domestik Regional Bruto (Safitri et al, 2022), upah rata-rata (Sari, 2021), tingkat

pendidikan (Berliani, 2021), jumlah merokok (Almizi dan Hermawati, 2018), dan akses seluler (Rini dan Sugiharti, 2023).

Terdapat beberapa variabel bebas yang digunakan dalam penelitian ini yang dapat menyebabkan terjadinya multikolinear. Yang dimaksud dengan multikolinear adalah keadaan dimana terdapat tingkat korelasi tinggi antara variabel bebas dalam model regresi (Dewanata, 2022). Adanya multikolinear dapat menyebabkan kesimpulan yang salah. Oleh karena itu, diperlukan suatu metode statistic yang dapat mengatasi terjadinya multikolinear. Metode tersebut adalah regresi ridge (Ameliana, 2023)

Dengan demikian tujuan dari penelitian ini adalah menentukan faktor yang berpengaruh terhadap kemiskinan dengan menggunakan regresi ridge. Adapun hasil dari penelitian ini diharapkan dapat memberikan sumbangan pengetahuan dan informasi yang dapat membantu pihak tertentu dalam mengatasi terjadinya kemiskinan.

## Metode

### A. Data

Penelitian ini menggunakan data sekunder tahun 2022 yang diperoleh dari Badan Pusat Statistik (BPS). Ada dua jenis variabel yang digunakan dalam penelitian ini, yaitu variabel bebas ( $X$ ) dan variabel terikat ( $Y$ ). Berikut adalah penjelasan singkat dari variabel penelitian yang digunakan pada Tabel 1.

*Tabel 1. Variabel Penelitian*

Variabel	Simbol	Keterangan
IPM	$X_1$	Capaian Pembangunan Manusia berbasis dari sejumlah komponen dasar kualitas hidup.
TPT	$X_2$	Persentase jumlah penggangguran terhadap jumlah angkatan kerja.
PDRB	$X_3$	Jumlah nilai tambah produksi dari berbagai sektor atau lapangan usaha di suatu daerah atau regional (Lube, Kalangi, dan Tolosang, 2021.).
Upah Rata-Rata	$X_4$	Rata-rata upah yang didapatkan oleh penduduk Indonesia yang dihitung tiap provinsinya
Lulus SD	$X_5$	Persentase penduduk Indonesia yang menyelesaikan pendidikan Sekolah Dasar (SD).

Lulus SMP	X6	Persentase penduduk Indonesia yang menyelesaikan Pendidikan Sekolah Menengah Pertama (SMP).
Lulus SMA	X7	Persentase penduduk Indonesia yang menyelesaikan Pendidikan Sekolah Menengah Atas (SMA).
Lama Sekolah	X8	Perhitungan lamanya sekolah penduduk Indonesia. (Kevin,Bhinadi dan Syari'udin, 2022).
Melek Huruf	X9	Perbandingan dari jumlah penduduk yang sudah dapat membaca dan juga menulis pada usia 15 tahun ke atas (Megantara, Kembar dan Budhi, 2020.)
Merokok	X10	Persentase dari banyaknya penduduk Indonesia yang merokok secara aktif
Seluler	X11	Persentase dari banyaknya penduduk Indonesia yang menggunakan telepon seluler secara aktif.
Persentase Kemiskinan	Y	Persentase penduduk yang berada dibawah Garis Kemiskinan (GK).

## B. Tahapan Penelitian

Penelitian ini dilakukan dengan tahapan sebagai berikut. Pertama, melakukan riset eksploratori. Kedua, pengumpulan data. Ketiga, melakukan analisis deskriptif. Keempat, pembuatan model regresi linear ganda. Kelima, pengecekan asumsi residual dan multikolinear. Keenam, pembuatan model regresi ridge. Ketujuh, penentuan variabel bebas yang signifikan dengan menggunakan uji t. Kedelapan, penentuan model terbaik dengan menggunakan AIC dan BIC. Kesembilan, analisis hasil yang diperoleh yang dilanjutkan dengan pembuatan kesimpulan. Kesepuluh, pembuatan artikel ilmiah.

## C. Model Regresi Linear

Regresi linear adalah metode statistik yang digunakan untuk memahami hubungan antara dua variabel atau lebih. Regresi linear berfungsi sebagai teknik statistik yang digunakan dengan tujuan membuat konsep korelasi antara dua variabel atau lebih. Model regresi linear digunakan untuk memprediksi hubungan antara variabel bebas dan variabel terikat dengan mengasumsikan adanya hubungan linear. Secara umum persamaan regresi linear diberikan pada Persamaan (1).

$$Y = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (1).$$

Keterangan:

$Y$  = Nilai dari variabel terikat pada observasi.

$\beta_p$  = Parameter model pada observasi ke- $p$ .

$X_p$  = nilai variabel bebas  $X$  pada observasi ke- $p$ .

$\varepsilon$  = Komponen error (pengaruh variabel bebas lain selain variabel  $X$ ).

#### D. Uji Residual

Model regresi linear yang baik adalah model yang memenuhi beberapa asumsi, asumsi tersebut adalah asumsi residual yang terdiri atas asumsi normalitas, asumsi homoskedastisitas, dan asumsi nonautokorelasi. Selain itu asumsi nonmultikolinear juga harus terpenuhi.

##### 1. Asumsi Normalitas

Asumsi normalitas adalah asumsi yang harus dipenuhi oleh model regresi, dimana error harus berdistribusi normal. Salah satu uji statistik yang dapat digunakan adalah uji Kolgomorov-Smirnov. Adapun hipotesis yang digunakan adalah:

$H_0$  : Error berdistribusi normal.

$H_1$  : Error tidak berdistribusi normal.

Tolak  $H_0$  jika  $p\text{-value} < 0.05$ .

##### 2. Asumsi Homoskedastisitas

Asumsi Homoskedastisitas merupakan asumsi yang harus terpenuhi oleh model regresi, pada asumsi ini error harus memiliki variansi yang konstan. Pelanggaran dari homoskedastisitas disebut dengan heteroskedastisitas, yaitu nilai variansi error berarti tidak konstan (Saputra dan Basuki, 2024). Adapun hipotesis yang digunakan adalah:

$H_0$  : Varians dari error bersifat konstan (homoskedastisitas).

$H_1$  : Varians dari error bersifat tidak konstan (heteroskedastisitas).

Tolak  $H_0$  jika  $p\text{-value} < 0.05$

##### 3. Asumsi Non Autokorelasi

Uji autokorelasi dalam analisis regresi linier menilai apakah error model regresi menunjukkan *independensi* satu sama lain. Ketika nilai variabel terikat pada suatu waktu tidak *independen* terhadap nilainya pada titik sebelumnya, maka terjadi autokorelasi. Salah satu uji statistik yang dapat digunakan adalah *Durbin-Watson*. Adapun hipotesis yang digunakan adalah:

$H_0$  : Error tidak terjadi autokorelasi.

$H_1$  : Error berautokorelasi.

Tolak  $H_0$  jika  $p\text{-value} < 0.05$

### E. Uji Multikolinearitas

Analisis regresi menggunakan uji multikolinearitas untuk mengetahui apakah dua atau lebih prediktor dalam suatu model mempunyai tingkat korelasi sedang atau tinggi satu sama lain. Adanya multikolinearitas dalam suatu penelitian dapat menimbulkan banyak permasalahan, antara lain kesalahan standar yang berlebihan sehingga dapat mempersulit proses identifikasi dampak spesifik setiap variabel bebas terhadap variabel terikat.

Untuk mengetahui terjadinya multikolinear dapat menggunakan *Variance Inflation Factor* (VIF). Terjadi multikolinear jika  $VIF > 10$ . Untuk menentukan nilai VIF dapat menggunakan Persamaan (2).

Rumus VIF adalah :

$$VIF = \frac{1}{1-R_j^2} \quad (2).$$

Keterangan :

$R_j^2$  = Koefisien determinasi dari regresi variabel bebas ke-j.

Hipotesis dari asumsi multikolinearitas adalah :

$H_0$ : Tidak terjadi multikolienaritas antara variabel bebas.

$H_1$ : Setidaknya ada satu multikolienaritas yang terjadi diantara variabel bebas.

### F. Uji F dan Uji t

Uji F adalah uji yang digunakan untuk menentukan pengaruh variabel bebas secara serentak. Uji F berfungsi untuk mengidentifikasi model linear yang paling sesuai untuk populasi dari mana data diambil. Hipotesis yang digunakan adalah:

$H_0$ : Tidak ada variabel bebas yang berpengaruh signifikan.

$H_1$ : Ada setidaknya satu variabel bebas yang berpengaruh signifikan.

Tolak  $H_0$  jika p-value < 0.05

Uji t adalah uji yang digunakan untuk menentukan variabel bebas secara parsial.

Uji ini menentukan variabel yang berpengaruh secara satu per satu. Sedangkan, hipotesis yang digunakan dalam uji t adalah:

$H_0$ : Variabel bebas tidak berpengaruh signifikan pada variabel terikat.

$H_1$ : Variabel bebas berpengaruh signifikan pada variabel terikat.

### G. Model Regresi Ridge

Model regresi ridge berfungsi sebagai alat untuk menangani tingkat korelasi yang tinggi antar variabel bebas. Model ini pertamakali diperkenalkan oleh Hoerl dan Kennard pada tahun 1970.

Secara umum model regresi ridge sama seperti model regresi linear yang dituliskan pada Persamaan (1). Perbedaannya terdapat pada cara memperoleh nilai taksiran. Nilai taksiran parameter ridge diperoleh dengan menggunakan Persamaan (3).

$$\hat{\beta}_{RR} = (X^T X + \lambda I)^{-1} X^T Y \quad (3).$$

Keterangan :

$\hat{\beta}_{RR}$  = Penduga parameter regresi *ridge-robust*.

X = Matriks desain yang berisi nilai dari semua variabel bebas.

Y = Vektor target.

$\lambda$  = Parameter penalty yang mengontrol kekuatan regularisasi.

I = Matriks identitas.

### H. R-Squared

Salah satu metode statistik yang dapat digunakan untuk menilai kinerja model regresi adalah *R-squared* ( $R^2$ ). Penilaian kesesuaian variabel bebas dalam memprediksi variabel terikat dapat diperoleh dari nilai *r-squared*. Nilai  $R^2$  yang rendah menunjukkan adanya keterbatasan yang signifikan kemampuan variabel bebas dalam menjelaskan perubahan variabel terikat (Jufrizen dan Sitorus 2021).

Adapun rumus untuk mencari  $R^2$  ditulis dalam Persamaan (4).

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{SSE_{error}}{SST_{total}} \quad (4).$$

### I. Akaike Information Criterion (AIC)

AIC adalah metode analisis yang biasa digunakan untuk mendapatkan model faktor produksi terbaik dengan menggunakan estimasi kemungkinan maksimum sebagai perhitungan yang tepat. Model dengan AIC yang lebih rendah merupakan model yang lebih baik karena memberikan hasil yang lebih akurat dan efisien.

### J. Bayesian Information Criterion (BIC)

Untuk memilih model yang paling sesuai dengan data, salah satu teknik evaluasi model adalah BIC. Dengan menghilangkan variabel-variabel yang tidak penting dari model, BIC menjadi alternatif bagi AIC. Model yang lebih baik adalah model yang BIC-nya lebih tinggi.

## Hasil dan Pembahasan

### A. Analisis Deskriptif

Sebelum membuat model regresi, terlebih dahulu dilakukan analisis deskriptif. Untuk statistic deskriptif diberikan pada Tabel 2.

*Tabel 2. Statistik Deskriptif Variabel Terikat*

<i>Min</i>	<i>1Q</i>	<i>Mean</i>	<i>3Q</i>	<i>Max</i>
4.530	6.338	10.251	12.120	26.680

Berdasarkan Tabel 2, persentase kemiskinan berkisar antara 4.53 hingga 26.68. Dengan nilai rata-rata 10.251 dan nilai kuartil pertama 6.338, serta 12.120 sebagai kuartil ketiga. Provinsi dengan persentase kemiskinan tertinggi secara berurut adalah Papua, Papua Barat, Nusa Tenggara Timur, Maluku, dan Gorontalo.

### B. Model Regresi Linear

Setelah kita membuat analisis deskriptif, kita telah mendapatkan model regresi linear yang dapat dilihat pada Tabel 3.

*Tabel 3. Koefisien Model Regresi Linear*

<i>Variable</i>	<i>Coefficient</i>	<i>p-value(t)</i>	<i>p-value(F)</i>
<b>Intercept</b>	61.1973	0.11653	
<b>X<sub>1</sub></b>	-0.4429	0.06642	
<b>X<sub>2</sub></b>	-1.0879	0.01452	
<b>X<sub>3</sub></b>	-0.5458	0.00034	
<b>X<sub>4</sub></b>	-0.0002	0.21368	
<b>X<sub>5</sub></b>	-0.1921	0.69604	1.312 x 10 <sup>-7</sup>
<b>X<sub>6</sub></b>	0.3241	0.11268	
<b>X<sub>7</sub></b>	-0.1939	0.04473	
<b>X<sub>8</sub></b>	5.3165	0.00059	
<b>X<sub>9</sub></b>	-0.0762	0.84443	
<b>X<sub>10</sub></b>	0.3968	0.002870	
<b>X<sub>11</sub></b>	-0.6055	0.00510	



Dari hasil model regresi linear yang telah dibentuk, diketahui bahwa variabel bebas yang digunakan dapat menjelaskan 80,41% dari variabel terikat yaitu persentase kemiskinan. Dengan hasil model regresi pada Persamaan (5).

$$\hat{Y} = 61.197 - 0.443X_1 - 1.088X_2 - 0.546X_3 - 0.0002X_4 - 0.192X_5 + 0.324X_6 - 0.091X_7 + 5.316X_8 - 0.076X_9 + 0.397X_{10} - 0.605X_{11} \quad (5).$$

### C. Uji Asumsi

Setelah memperoleh model regresi linear, dilakukan uji asumsi untuk memastikan bahwa model regresi linear memenuhi seluruh uji asumsi, diantaranya berdistribusi normal, memiliki varians konstan, tidak berautokorelasi, dan tidak terdapat multikolinearitas dalam model yang telah dibuat.

*Tabel 4. Hasil Uji Asumsi*

Uji Asumsi	Keputusan	Keterangan
Uji Normalitas	0.114	Gagal tolak H0, error berdistribusi normal
Uji Homoskedastisitas	0.8947	Gagal Tolak H0, error memiliki varians konstan
Uji Autokorelasi	0.06371	Gagal Tolak H0, error tidak berautokorelasi
Uji Multikolinearitas	>10	Terdapat multikolinearitas

Berdasarkan Tabel 4, dapat disimpulkan bahwa model yang telah dibuat berdistribusi normal, memiliki varians konstan (homoskedastis), dan tidak berautokorelasi, namun, ditemukan adanya multikolinearitas. Untuk selanjutnya dibuat model regresi ridge untuk mengatasi pengaruh multikolinearitas ini.

### D. Model Regresi Ridge

Regresi ridge, sebuah model regresi yang menambahkan parameter c sebagai konstanta bias pada matriks varians-kovarians variabel terikat, merupakan salah satu metode untuk mengatasi multikolinearitas. Secara lengkap untuk nilai koefisien ridge beserta uji F dan uji t diberikan pada Tabel 5.

*Tabel 5. Koefisien Model Regresi Ridge*

Variable	Coefficient (Linear Regression)	Coefficient (Ridge Regression)	Standard Error	t-statistic	p-value
Intercept	61.1973	108278.7098	80561.1427	1.3441	0.1920
$X_1$	-0.4429	-9.9336	5.0465	-1.9684	0.0612

$X_2$	-1.0879	-9.6418	3.5705	-2.7004	0.0128
$X_3$	-0.5458	-11.6139	2.7058	-4.2922	0.0003
$X_4$	-0.0002	-5.8815	4.5020	-1.3064	0.2043
$X_5$	-0.1921	-3.4131	8.4459	-0.4041	0.6899
$X_6$	0.3241	11.0774	6.5750	1.6848	0.1056
$X_7$	-0.1939	-12.1018	5.5801	-2.1687	0.0407
$X_8$	5.3165	24.9783	6.1430	4.0661	0.0005
$X_9$	-0.0762	-1.4273	7.0403	-0.2027	0.8411
$X_{10}$	0.3968	8.5330	3.5790	2.3842	0.0258
$X_{11}$	-0.6055	-25.9585	8.2097	-3.1620	0.0044

  

<i>F-statistic</i>	<i>R-squared</i>	<i>Adjusted R-squared</i>	<i>Model p-value</i>
14.28	0.8675	0.8123	$5.57 \times 10^{-8}$

Dari hasil model regresi ridge yang telah dibentuk, diketahui bahwa variabel bebas yang digunakan dapat menjelaskan 81,23% dari variabel terikat yaitu persentase kemiskinan. Dengan hasil model regresi pada Persamaan (6).

$$\hat{Y} = 108278.7098 - 9.934X_1 - 9.642X_2 - 11.614X_3 - 5.882X_4 - 3.413X_5 + 11.077X_6 - 12.102X_7 + 24.978X_8 - 1.427X_9 + 8.533X_{10} - 25.959X_{11} \quad (6).$$

#### E. Uji F dan Uji t

Pengujian signifikansi variabel dalam penelitian ini dilakukan dengan membandingkan nilai *P-value* yang diperoleh dari model regresi linear yang telah dibuat. Nilai *P-value* yang telah diperoleh dari uji F pada Tabel 5 menunjukkan nilai  $1,312e^{-07}$ , yang mengindikasikan adanya penolakan terhadap hipotesis nol, dikarenakan nilainya yang lebih kecil dibandingkan *level of significant* (0,05). Oleh karena itu, dapat disimpulkan bahwa terdapat setidaknya satu variabel bebas yang menunjukkan korelasi yang signifikan secara statistik dengan variabel terikat. Untuk mengetahui variabel bebas yang memiliki hubungan signifikan secara statistik dengan variabel terikat, maka dilakukan uji signifikansi parsial.

Dari nilai *P-value* setiap variabel bebas pada Tabel 5, didapatkan TPT, PDRB, Lulus SMA, Lama Sekolah, Rokok, dan Seluler memiliki pengaruh yang signifikan

terhadap variabel terikat yaitu persentase kemiskinan. Kesimpulan ini diperoleh dari nilai *P-value* keenam variabel yang lebih kecil dari *level of significant* (0,05), yang menyebabkan ditolaknya  $H_0$ . Keenam variabel tersebut dapat dijadikan fokus utama dari pihak berwenang dalam usaha menurunkan persentase kemiskinan di Indonesia.

#### F. Model Ridge Signifikan

Tabel 6. Koefisien Model Regresi Ridge Signifikan

<i>Variable</i>	<i>Coefficient (Linear Regression)</i>	<i>Coefficient (Ridge Regression)</i>	<i>Standard Error</i>	<i>t-statistic</i>	<i>p-value</i>
Intercept	26.1083	2879.3303	436.0372	6.6034	$<2e^{-16}$
X2	-1.4680	-13.0097	3.4493	-3.7716	0.0008
X3	-0.4858	-10.3368	2.7787	-3.7201	0.0009
X7	-0.2290	-14.2871	4.0798	-3.5019	0.0016
X8	5.3598	25.1816	4.8955	5.1438	$<2e^{-16}$
X10	0.5624	12.0943	3.1808	3.8022	0.0007
X11	-0.5976	-25.6197	3.5847	-7.1470	$<2e^{-16}$

  

<i>F-statistic</i>	<i>R-squared</i>	<i>Adjusted R-squared</i>	<i>Model p-value</i>
17.9997	0.7883	0.7518	1.3728e-08

Dari uji signifikansi yang telah dilakukan, terdapat 6 variabel yang berpengaruh signifikan terhadap variabel terikat. Dengan variabel tersebut dibuatlah model regresi ridge signifikan, diketahui bahwa variabel bebas yang digunakan dapat menjelaskan 75.18%% dari variabel terikat yaitu persentase kemiskinan. Dengan hasil model regresi pada Persamaan (7).

$$\hat{Y} = 2879.3303 - 13.0097X_2 - 10.3368X_3 - 14.2871X_7 + 25.1816X_8 + 12.0943X_{10} - 25.6197X_{11} \quad (7)$$

Konstan sebesar 2879,3303 menunjukkan nilai harapan dari persentase kemiskinan ketika semua variabel bebas bernilai 0. Persentase kemiskinan mengalami penurunan sebesar 13,0097 ketika terjadi kenaikan persentase jumlah pengangguran terhadap jumlah angkatan kerja sebesar satu-satuan dan variabel bebas lainnya tetap. Persentase kemiskinan mengalami penurunan sebesar 10,3368 ketika terjadi kenaikan jumlah nilai tambah produksi dari berbagai sektor atau

lapangan usaha di suatu daerah atau regional sebesar satu-satuan dan variabel bebas lainnya tetap. Persentase kemiskinan mengalami penurunan sebesar 14,2871 ketika terjadi kenaikan persentase penduduk Indonesia yang menyelesaikan Pendidikan Sekolah Menengah Atas (SMA) sebesar satu-satuan dan variabel bebas lainnya tetap. Persentase kemiskinan mengalami kenaikan sebesar 25,1816 ketika terjadi kenaikan perhitungan lamanya sekolah penduduk Indonesia sebesar satu-satuan dan variabel bebas lainnya tetap. Persentase kemiskinan mengalami kenaikan sebesar 12,0943 ketika terjadi kenaikan persentase dari banyaknya penduduk Indonesia yang merokok secara aktif menunjukkan sebesar satu-satuan dan variabel bebas lainnya tetap. Persentase kemiskinan mengalami penurunan sebesar 25,6197 ketika terjadi kenaikan persentase dari banyaknya penduduk Indonesia yang menggunakan telepon seluler secara aktif sebesar satu-satuan dan variabel bebas lainnya tetap.

### G. Perbandingan Metrik

Karena nilai *R-Squared* dari kedua model, model linear dan ridge, tidak memiliki perbedaan yang signifikan. Metrik yang digunakan untuk membandingkan kedua model adalah AIC dan BIC. Nilai metrik tersebut tertera pada Tabel 7.

Tabel 7. Metrik Perbandingan

<i>Metric</i>	<i>Linear Regression</i>	<i>Significant Ridge Regression</i>
AIC	168.8447	71.9133
BIC	189.0642	205.6826

Dari data perbandingan metrik di atas, dapat disimpulkan bahwa model regresi ridge signifikan memiliki performa yang lebih baik dibandingkan dengan model regresi linear. Kesimpulan ini dapat dibuat dengan melihat analisis komparatif dari data yang tersedia, yang menunjukkan nilai AIC yang jauh lebih rendah untuk model regresi ridge signifikan ketika disandingkan dengan model regresi linear. Model regresi ridge signifikan menunjukkan kinerja yang lebih baik, seperti yang ditunjukkan oleh nilai AIC yang lebih rendah yaitu 71.9133 dibandingkan dengan nilai AIC model regresi linear sebesar 168,8447. Selain itu, nilai BIC juga mendukung kesimpulan ini, dengan model regresi linear mencatat nilai BIC sebesar 189,0642 dan model regresi ridge signifikan memiliki nilai BIC yang sedikit lebih tinggi, yaitu 205.6826. Oleh karena itu, perbandingan nilai AIC dan BIC menunjukkan bahwa model regresi ridge mengungguli model regresi linear dalam analisis ini.

## Kesimpulan

Dengan menggunakan metode analisis regresi, beberapa faktor dalam penelitian ini yang memiliki pengaruh signifikan terhadap persentase kemiskinan di Indonesia dapat disimpulkan menjadi enam faktor yaitu tingkat pengangguran terbuka(TPT), produk domestik regional bruto (PDRB), jumlah lulusan sekolah menengah atas (SMA), lama sekolah, jumlah perokok, dan ketersediaan telepon seluler. Pada penelitian ini, model regresi linear terbukti tidak efektif, dimana model tersebut memiliki masalah multikolinearitas yang dapat menyebabkan hasil penelitian tidak dapat diandalkan. Untuk mengatasi adanya multikolinear antar variabel bebas, dapat digunakan adalah model regresi ridge. Metode ini terbukti lebih unggul dibandingkan dengan metode regresi linear karena dapat mengatasi kasus multikolinear dengan hasil yang lebih baik berdasarkan hasil perbandingan metrik antara model regresi linear dan regresi ridge yang telah dilakukan.

## Ucapan Terima Kasih

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena atas rahmat-Nya, tim penulis dapat menyelesaikan makalah tentang “Pemodelan Persentase Kemiskinan Di Indonesia Dengan Menggunakan Metode Regresi Ridge” hingga selesai. Penulis ingin berterima kasih kepada:

1. Ibu Margaretha Ohyver, S.Si., M.Si selaku dosen pendamping, atas bimbingan dan dukungannya.
2. Rekan-rekan yang selalu mendukung penulis.
3. Penulis juga mengucapkan terima kasih kepada pihak-pihak yang telah membantu dan menjadi sumber informasi selama proses pengerjaan makalah ini.

Penulis menyadari bahwa banyak kendala dalam penulisan karya ini. Namun berkat rahmat Tuhan Yang Maha Esa dan dukungan berbagai pihak, kendala tersebut dapat diatasi dengan baik. Penulis juga ingin meminta maaf atas segala kesalahan dalam penulisan karya ini. Semoga makalah ini dapat bermanfaat bagi semua pihak.

## Kontribusi Penulis

Penulis Satu mengidentifikasi masalah dan menentukan topik yang dipilih, melakukan penelitian, dan terlibat dalam penulisan artikel ilmiah; Penulis Kedua berperan dalam melakukan pengumpulan data sekunder secara sistematis dan logis, melakukan penelitian, dan mengolah data sekunder, Melakukan penelitian, dan terlibat dalam penulisan artikel ilmiah.; Penulis Ketiga berperan dalam melakukan riset exploratory data sekunder yang didapatkan, melakukan penelitian, dan pengolahan serta interpretasi output penelitian dan terlibat dalam penulisan artikel

ilmiah.; Penulis Keempat melakukan arahan penelitian dan penyelesaian naskah/manuskrip.

### **Daftar Pustaka**

- Ameliana, W. (2023). Penanganan Masalah Multikolinearitas Dengan Regresi Ridge, Lasso, dan Elastic-Net Pada Kasus Balita Stunting di Indonesia.
- Bakhri, S. and Fauzi, P., 2021, March. Significant role of economic growth in alleviating poverty. In 1st Paris Van Java International Seminar on Health, Economics, Social Science and Humanities (PVJ-ISHESSH 2020) (pp. 504-507). Atlantis Press.
- Berliani, K., 2021. Pengaruh Tingkat Pengangguran, Tingkat Pendidikan dan Laju Pertumbuhan Penduduk Terhadap Tingkat Kemiskinan Penduduk Provinsi Jawa Barat Tahun 2015-2020. *Syntax Literate; Jurnal Ilmiah Indonesia*, 6(2), p.872.
- Dewanata, P., 2022. Analisis Volume Saham Campina Ice Cream Industry Tbk.(ISSI): Regresi dengan Teknik Imputasi LOCF: Regresi dengan Teknik Imputasi LOCF. *Indonesian Journal of Islamic Economics and Business*, 7(2), pp.72-93..
- Jufrizen, J. and Sitorus, T.S., 2021, July. Pengaruh Motivasi Kerja dan Kepuasan Kerja Terhadap Kinerja Dengan Disiplin Kerja Sebagai Variabel Intervening. In *Seminar Nasional Teknologi Edukasi Sosial dan Humaniora* (Vol. 1, No. 1, pp. 844-859).
- Kevin, A.V., Bhinadi, A. and Syari'udin, A., 2022. Pengaruh PDRB, Angka Harapan Hidup, dan Rata Rata Lama Sekolah Terhadap Kemiskinan di Kabupaten/Kota Provinsi Jawa Tengah Tahun 2013-2021. *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, 1(12), pp.2959-2968.
- Lube, F., Kalangi, J.B. and Tolosang, K.D., 2021. Analisis Pengaruh Upah Minimum dan PDRB terhadap Penyerapan Tenaga Kerja di Kota Bitung. *Jurnal Berkala Ilmiah Efisiensi*, 21(3).
- Megantara, D.E., Kembar, M. and Budhi, S., 2020. Pengaruh Angka Melek Huruf Dan Upah Minimum Terhadap Tingkat Pengangguran Dan Indeks Pembangunan Manusia Kabupaten/Kota Di Provinsi Bali. *Jurnal Ekonomi Pembangunan Universitas Udayana*, 9(1), pp.91-119.
- Prasada, I.Y., Yulhar, T.F.M. and Rosa, T.A., 2020. Determinants of poverty rate in Java Island: Poverty alleviation policy. *Jurnal Ekonomi Pembangunan*, 18(2), pp.95-104.
- Ristika, E.D., Primandhana, W.P. and Wahed, M., 2021. Analisis Pengaruh Jumlah Penduduk, Tingkat Pengangguran Terbuka Dan Indeks Pembangunan Manusia Terhadap Tingkat Kemiskinan Di Provinsi Jawa Timur. *Eksis: Jurnal Ilmiah Ekonomi Dan Bisnis*, 12(2), pp.129-136.

- Saputra, M.R.H., Basuki, R.R. and Muhtadin, I.A., 2024. Analisis Regresi Pada Pelanggaran Asumsi Klasik Pada Regresi Linear. Madani: Jurnal Ilmiah Multidisiplin, 2(1).
- Sari, Y.A., 2021. Pengaruh Upah Minimum, Tingkat Pengangguran Terbuka Dan Jumlah Penduduk Terhadap Kemiskinan Di Provinsi Jawa Tengah. Equilibrium, 10(2).