Research Review
Jurnal Ilmiah Multidisiplin
Vol. 4, No. 1, Juni 2025 ISSN: 29622743

Perbandingan *Jackknife Ridge Regression* dan *Principal Component Regression* dalam Penanganan Kasus Multikolinearitas (Studi Kasus: Indeks Pembangunan Manusia di Indonesia)

A Comparison between Jackknife Ridge Regression and Principal Component Regression in Handling Multicollinearity Cases (Case Study : Human Development Index in Indonesia)

Nur'ain Manoppo[1], La Ode Nashar[2], Djihad Wungguli[3], Muhammad Rezky F. Payu[4], Siti Nurmardia Abdussamad[5], Salmun K. Nasib[6]
[1,2,4,5,6]Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Gorontalo
[3]Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Gorontalo
nurainmnp@gmail.com[1], laode.nashar@ung.ac.id[2], djihad@ung.ac.id[3], rezky@ung.ac.id[4]
sitinurmardia@ung.ac.id[5], salmun@ung.ac.id[6]

Article Info

Article history:

Received: 10 Juni 2025
Revised: 26 Juni 2025
Accepted: 27 Juni 2025

Keywords:

Human Development Index
Multicollinearity
Jackknife Ridge Regression
Principal Component Regression

Kata Kunci:
Indeks Pembangunan Manusia
Multikolinearitas
Regresi Ridge Jackknife
Regresi Komponen Utama

Abstract

According to data from Statistics Indonesia, the Human Development Index (HDI) in 2022 reached 72.91, increasing from 72.29 in the previous year. Although Indonesia's HDI continues to improve, disparities remain among provinces, indicating that HDI distribution is still uneven. Given the importance of HDI in aregion, it is necessary to conduct statistical analysis to identify the factors that significantly influence HDI using regression analysis. In applying multiple linear regression, several classical statistical assumptions must be met, one of which is the central focus of this analysis-addressing the issue of multicollinearity. Several methods have been identified to address multicollinearity, including Jackknife Ridge Regreesion (JRR) and Principal Component Regression (PCR). This study aims to compare the effectiveness of both methods in handling multicollinearity based on Adjusted R^2 and Mean Square Error (MSE) and to analyze the factors that significantly influence the HDI level in Indonesia. The data used in this study are secondary data comprising HDI and its related factors for each province in Indonesia in 2022, obtained from bps.go.id. Based on the analysis, the best model uses the JRR method, with an Adjusted R^2 value of 96.7% and MSE of 0.033.

Abstrak

Menurut data dari Badan Pusat Statistik Indonesia, Indeks Pembangunan Manusia (IPM) pada tahun 2022 mencapai 72,91 yang menunjukkan peningkatan dari 72,29 pada tahun sebelumnya. Meskipun IPM di Indonesia terus meningkat, masih terdapat perbedaan di tiap provinsi yang menunjukkan bahwa distribusi IPM di Indonesia belum merata. Melihat pentingnya IPM pada suatu wilayah, maka perlu dilakukan analisis secara statistik dengan mengidentifikasi faktor-faktor yang berpengaruh signifikan terhadap IPM dengan menggunakan analisis regresi. Pada penerapan regresi linier berganda memerlukan pemenuhan beberapa asumsi klasik secara statistik yang salah satunya menjadi titik pusat analisis yaitu penanganan masalah multikolinearitas. Beberapa metode telah diidentifikasi dapat

mengatasi masalah multikolinearitas antara lain *Jackknife Ridge Regression* (JRR) dan *Principal Component Regression* (PCR). Tujuan dari penelitian ini yakni untuk membandingkan efektivitas antara kedua metode dalam mengatasi multikolinearitas berdasarkan nilai *Adjusted* R^2 dan *Mean Square Error* (MSE), serta menganalisis faktor-faktor yang memberikan pengaruh signifikan terhadap tingkat IPM di Indonesia. Dalam penelitian ini data yang digunakan adalah data sekunder yang mencakup data IPM serta faktor yang terkait pada tiap provinsi di Indonesia tahun 2022, yang diperoleh melalui laman *bps.go.id*. Berdasarkan hasil analisis yang telah dilakukan, diperoleh model terbaik yaitu model dengan metode JRR dengan nilai *Adjusted* R^2 sebesar 96,7% dan MSE sebesar 0,033.

Corresponding Author:

Nur'ain Manoppo
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Gorontalo
nurainmnp@gmail.com

1. PENDAHULUAN

Penyempurnaan kualitas Sumber Daya Manusia (SDM) diakui menjadi indikator keberhasilan utama dalam pembangunan ekonomi, sebagaimana dijelaskan oleh (Si'Iang et al., 2019). Sumber daya manusia harus memiliki kemampuan untuk berkembang dan mengoptimalkan kemampuan yang dimiliki untuk mencapai pembangunan yang berkelanjutan. Pembangunan dapat didefinisikan secara sederhana sebagai upaya ataupun proses untuk mengubah sesuatu menjadi lebih baik dari sebelumnya (Sangkereng, Engka, & Sumual, 2019). Salah satu indikator yang digunakan untuk mengukur SDM adalah Indeks Pembangunan Manusia (IPM) (Ningrum et al., 2020). Menurut data dari Badan Pusat Statistik (BPS) Indonesia, IPM pada tahun 2022 mencapai 72,91, angka ini menunjukkan peningkatan dari 72,29 pada tahun sebelumnya. Meskipun IPM Indonesia terus meningkat, masih terdapat perbedaan di tiap provinsi, yang menunjukkan bahwa distribusi IPM di Indonesia belum merata. Melihat pentingnya IPM pada suatu wilayah, maka perlu dilakukannya analisis mengenai IPM (Arum & Fitri, 2023). Analisis dapat dilakukan secara statistik dengan mengidentifikasi faktor-faktor yang berpengaruh signifikan terhadap IPM melalui metode analisis regresi.

Analisis regresi adalah salah satu teknik statistik yang paling umum digunakan yang bertujuan untuk menentukan hubungan sebab akibat dari satu atau lebih variabel independen terhadap variabel dependen (Taqiyyuddin & Irfan, 2022). Pada penerapan regresi linier berganda, memerlukan pemenuhan beberapa asumsi klasik secara statistik yang diantaranya, normalitas, autokorelasi, heteroskedastisitas, serta multikolinearitas. Penanganan masalah multikolinearitas menjadi titik pusat analisis ini karena, asumsi multikolinearitas seringkali men jadi masalah dalam analisis regresi karena gagal dipenuhi yang akan menghasilkan nilai error yang lebih besar dan berdampak pada hasil keputusan yang kurang signifikan, seperti yang dijelaskan oleh Silitonga et al. (2021) dalam studinya. Hasil yang tidak signifikan sering menunjukkan bahwa model regresi gagal menjelaskan hubungan antara variabel dependen dan independen dengan tepat. Dalam mengatasi kasus multikolinearitas, beberapa metode telah diidentifikasi, antara lain *Jackknife Ridge Regression* (JRR) dan *Principal Component Regression* (PCR).

Beberapa peneliti telah melakukan penelitian mengena JRR dan juga mengenai PCR. Peneliti (Fatihah, 2022) memfokuskan penelitiannya pada penerapan metode JRR untuk mengatasi multikolinearitas. Penelitian yang dilakukan oleh Arrasyid et al. (2021) membandingkan *Generalized Ridge Regression* (GRR), JRR dan *Modified Jackknife Ridge Regrssion* (MJRR) untuk penanganan multikolinearitas. Hasil penelitiannya menyatakan bahwa ketiga metode tersebut dapat mengatasi multikolinearitas, namun dari ketiga metode tesebut metode JRR adalah yang paling baik karena memiliki nilai MSE terkecil. Sementara itu, (Maulida, 2022) memfokuskan penelitiannya pada perbandingan antara PCR dan *ridge regression* untuk menganalisis faktor-faktor yang berpengaruh terhadap IPM. Hasil penelitian menunjukkan bahwa faktor-faktor yang berpengaruh terhadap IPM berdasarkan model ridge regression yang merupakan model terbaik yaitu produk domestik regional bruto, angka harapan hidup dan rata-rata lama sekolah. (Amelia & Putra, 2023) memfokuskan penelitiannya pada pembuktian metode PCR dalam mengatasi multikolinearitas.

Berdasarkan uraian yang telah dijelaskan bahwa metode *jackknife ridge regression* dan *principal component regreesion* dapat mengatasi multikolinearitas, maka dari itu peneliti tertarik untuk

membandingkan efektivitas antara metode *jackknife ridge regression* dan *principal component regression* dalam mengatasi multikolinearitas berdasarkan nilai *Adjusted* R^2 dan *Mean Square Error* (MSE) serta menganalisis faktor-faktor yang memberikan pengaruh signifikan terhadap tingkat IPM di Indonesia.

2. METODE PENELITIAN

2.1. Data

Penelitian ini menggunakan data sekunder yang merupakan data yang diperoleh secara tidak langsung. Dalam hal ini data yang digunakan adalah data indeks pembangunan manusia serta faktor-faktor yang mempengaruhinya di 34 Provinsi di Indonesia yang diperoleh melalui laman *bps.go.id*. Variabel yang digunakan pada penelitian ini diberikan pada Tabel 1.

Tabel 1. Daftar Variabel Penelitian

Variabel	Keterangan	Satuan
Y	Indeks Pembangunan Manusia (IPM)	Persen
X_1	Jumlah Penduduk	Jiwa
X_2	Angka Harapan Hidup (AHH)	Tahun
X_3	Jumlah Penduduk Miskin	Jiwa
X_4	Harapan Lama Sekolah (HLS)	Tahun
X_5	Pengeluaran perkapita	Rupiah
X_6	Tingkat Pengangguran Terbuka (TPT)	Persen
X_7	Produk Domestik Regional Bruto (PDRB)	Rupiah

2.2. Regresi Linier Berganda

Regresi linier berganda merupakan persamaan regresi yang melibatkan lebih dari satu variabel independen (X). Regresi linier berganda bertujuan untuk mengukur pengaruh signifikan antara variabel independen dan variabel dependen (Gujarati & Porter, 2009). Secara umum model regresi linier yang mengandung p-variabel bebas menurut Sembiring dalam Syarifuddin dan Saudi (2022) dapat dituliskan sebagai berikut.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \varepsilon_i \quad (1)$$

Regresi linier dapat ditaksir dengan menggunakan Ordinary Least Square (OLS). Penduga parameter β menggunakan metode OLS adalah sebagai berikut (Gujarati & Porter, 2009).

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (2)$$

2.3. Standarisasi Data

Proses standarisasi data bertujuan untuk menyamakan satuan setiap variabel yang diteliti. Karena perbedaan satuan ini, maka akan menghasilkan rentang nilai yang luas antar variabel, sehingga akan mempengaruhi hasil keputusan analisis (Auqino et al., 2019). Oleh karena itu sebelum melakukan analisis dilakukan standarisasi data dengan persamaan berikut (Bau et al., 2023).

$$Z_{ip} = \frac{X_{ip} - \bar{X}_p}{S_p} \quad (3)$$

2.4. Korelasi Pearson

Korelasi Pearson digunakan untuk menentukan arah hubungan, kekuatan hubungan, dan signifikansi kuat hubungan antara dua variabel (Rusli & Setyawan, 2022). Uji korelasi pearson dapat dihitung menggunakan rumus berikut, berdasarkan Hadi dalam Roflin dan Zulvia (2021).

$$r_{XY} = \frac{n\sum XY - (\sum X \sum Y)}{\sqrt{(n\sum X^2 - (\sum X)^2)}\sqrt{(n\sum Y^2 - (\sum Y)^2)}} \quad (4)$$

Hipotesis dari uji korelasi sebagai berikut (Khoiri, 2021):

- H_0: Tidak terdapat korelasi
- H_1: Terdapat korelasi

Dengan dasar pengambilan keputusan, jika nilai signifikansi < 0,05 maka tolak H0 yang artinya terdapat korelasi.

2.5. Uji Multikolinearitas

Dalam analisis regresi linier berganda, ukuran yang digunakan untuk mendeteksi adanya multikolinearitas adalah dengan menggunakan *Variance Inflation Factors* (VIF) (Pendi, 2021). Nilai VIF didefinisikan sebagai berikut (Gujarati & Porter, 2009).

$$VIF = \frac{1}{(1 - R^2)} \tag{5}$$

Hipotesis dari uji multikolinearitas sebagai berikut.

- H_0: Tidak terdapat multikolinearitas
- H_1: Terdapat multikolinearitas

Dengan dasar pengambilan keputusan, jika nilai VIF < 10, maka tidak mengalami masalah multikolinearitas sehingga terima H_0 (Rahmadeni & Yunita, 2019).

2.6. *Jackknife Ridge Regression*

Jackknife dapat dikembangkan pada metode *Generalized Ridge Regression* (GRR) dengan melakukan *resampling jackknife*, seperti yang diungkapkan oleh Arrasyid et al. (2021). GRR merupakan pengembangan dari metode *ridge regression* (Anggraini et al., 2019). Berdasarkan Hoerl dan Kennard dalam Anggraini et al. (2019), jika dalam metode *ridge regression* hanya satu nilai konstanta yang diperlukan, maka pada metode GRR memerlukan sebuah konstanta (*k)* untuk setiap variabel independennya. Adapun langkah-langkah yang dilakukan untuk memperoleh model *jackknife ridge regression*, yaitu:

1) Mencari nilai eigen dan vektor eigen dari matriks X^TX (Arrasyid et al., 2021). Persamaan untuk memperoleh nilai eigen $(\lambda_1, \lambda_2, \cdots, \lambda_p)$ sebagai berikut.

$$|C - \lambda I| = 0 \tag{6}$$

Kemudian untuk memperoleh vektor eigen *(v)* didapatkan melalui persamaan berikut.

$$(C - \lambda I)v = 0 \tag{7}$$

2) Menghitung $\hat{\gamma}_{OLS}$ model linier kanonik, yang digunakan untuk menaksir $\hat{\gamma}_{GRR}$ dengan hanya menambahkan konstanta bias *k* dari $\hat{\gamma}_{OLS}$, untuk memperoleh $\hat{\gamma}_{OLS}$ dapat menggunakan persamaan berikut.

$$\hat{\gamma}_{OLS} = (D)^{-1}U^TY \tag{8}$$

3) Mencari nilai MSE OLS $(\hat{\sigma}^2)$ dengan menggunakan persamaan berikut.

$$\hat{\sigma}^2 = \frac{ESS}{n - j - 1} \tag{9}$$

4) Mencari nilai konstanta bias *k* dengan iterasi HKB (Hoerl, Kennard, dan Balwin), adapun langkah-langkahnya sebagai berikut (Arrasyid et al., 2021):
   a) Mencari nilai *k* awal dengan cara berikut, berdasarkan Hoerl dan Kennard dalam Fatihah (2022).

$$k_p^0 = \frac{\hat{\sigma}^2_{OLS}}{\hat{\gamma}^2_{OLS_P}} \tag{10}$$

   b) Menghitung parameter awal $\hat{\gamma}_{GRR}$ dengan mensubstitusikan nilai *k* awal yang telah diperoleh.

$$\hat{\gamma}_{GRR} = (D + k)^{-1}U^TY \tag{11}$$

   c) Menghitung nilai $\hat{\beta}_{GRR}$

$$\hat{\beta}_{GRR} = Q\hat{\gamma}_{GRR} \tag{12}$$

   d) Menghitung nilai MSE $(\hat{\sigma}^2)$ baru dengan menggunakan parameter $\hat{\gamma}_{GRR}$.
   e) Mencari nilai *k* baru dengan cara berikut.

$$k_p^1 = \frac{\hat{\sigma}^2_{GRR}}{\hat{\gamma}^2_{GRR_P}} \tag{13}$$

5) Proses iterasi untuk pemilihan nilai *k* selanjutnya akan berhenti jika $\left|(\hat{\gamma}_{GRR}{}^T\hat{\gamma}_{GRR})^i(\hat{\gamma}_{GRR}{}^T\hat{\gamma}_{GRR})^{i-1}\right| \leq 0{,}01$
6) Mengestimasi parameter JRR, dengan menghitung $\hat{\gamma}_{JRR}$ terlebih dahulu menggunakan persamaan berikut dengan nilai *k* yang telah diperoleh pada proses iterasi sebelumnya

$$\hat{Y}_{JRR} = \left(I - (A^{-1}k)\right)^2 \hat{Y}_{OLS} \tag{14}$$

Kemudian nilai parameter $\hat{\beta}_{JRR}$ akan diperoleh dengan menggunakan persamaan berikut.

$$\hat{\beta}_{JRR} = Q\hat{Y}_{JRR} \tag{15}$$

7) Mengembalikan model *jackknife ridge regression* kedalam bentuk variabel asal dengan menggunakan persamaan berikut.

$$\beta_P = \left(\frac{S_Y}{S_X}\right)\beta_p^*$$
$$\beta_0 = \bar{Y} - \beta_1\bar{X}_1 - \beta_2\bar{X}_2 - \cdots - \beta_P\bar{X}_P \tag{16}$$

2.7. *Principal Component Regression*

PCR adalah jenis analisis regresi yang digunakan untuk menentukan hubungan antara variabel terikat dan sejumlah skor komponen utama sebagai variabel independen (Fauzan & Soehardjoepri, 2023). Menurut Maulida (2022) pembentukan PCR melalui principal component yang bergantung pada matriks kovarian (Σ) atau matriks korelasi (ρ). Matriks kovarian (Σ) diterapkan ketika semua variabel yang diamati memiliki satuan yang homogen dan matriks korelasi (ρ) diterapkan apabila satuan dari variabel-variabel tersebut berbeda. PCR merupakan metode yang didasarkan pada *Principal Component Analysis* (PCA) yang bertujuan mereduksi dimensi variabel yang saling berkorelasi tetapi tidak mengurangi karakteristik variabel tersebut, melalui transformasi ke variabel baru (komponen utama) (Amelia & Putra, 2023). Adapun langkah langkah yang dilakukan untuk memperoleh model PCR, yaitu:

1) Melakukan perhitungan korelasi dari data yang telah distandarisasi dengan menggunakan persamaan berikut (Amelia & Putra, 2023).

$$r_{mn} = \frac{x_{mn}}{\sqrt{x_{mm}}\sqrt{x_{nn}}} = \frac{\sum_{i=1}^{n}(x_{im} - \bar{x}_m)(x_{in} - \bar{x}_n)}{\sqrt{\sum_{i=1}^{n}(x_{im} - \bar{x}_m)^2}\sqrt{\sum_{i=1}^{n}(x_{in} - \bar{x}_n)^2}} \tag{17}$$

2) Melakukan perhitungan nilai eigen dan vektor eigen dari matriks korelasi (ρ).
3) Menentukan jumlah komponen utama yang dilakukan dengan melihat scree plot, dimana scree plot tersebut merupakan plot antara nilai eigen (λ_p) dan komponen utama (Siburian et al., 2019).
4) Membentuk persamaan komponen utama dengan nilai vektor eigen. Vektor eigen merupakan koefisien untuk setiap variabel yang digunakan dalam membentuk komponen utama atau *loadings* (Enzellina & Suhaedi, 2022).

$$\begin{aligned} W_1 &= a_{11}Z_1 + a_{12}Z_2 + \cdots + a_{1j}Z_j \\ W_2 &= a_{21}Z_1 + a_{22}Z_2 + \cdots + a_{2j}Z_j \\ &\vdots \\ W_j &= a_{j1}Z_1 + a_{j2}Z_2 + \cdots + a_{jj}Z_j \end{aligned} \tag{18}$$

5) Menentukan skor komponen utama yang terbentuk guna membentuk persamaan regresi komponen utama.
6) Meregresikan komponen utama terpilih dengan variabel dependen. Model ini disebut model principal component regression.

$$\ddot{Y}_{PCR} = m_0 + m_1W_1 + m_2W_2 + \cdots + m_pW_p + \varepsilon \tag{19}$$

7) Mengembalikan model PCR ke bentuk variabel baku dengan mensubstitusi persamaan (18) ke dalam persamaan (19), sehingga model akan menjadi:

$$\hat{Y}_{PCR} = \beta_0 + \beta_1Z_1 + \beta_2Z_2 + \cdots + \beta_JZ_J + \varepsilon \tag{20}$$

8) Mengembalikan model PCR ke dalam bentuk variabel asal

$$\begin{aligned} \hat{Y}_{PCR} &= \beta_0 + \beta_1\frac{(X_1 - \bar{X}_1)}{\sqrt{\sigma_1}} + \beta_2\frac{(X_2 - \bar{X}_2)}{\sqrt{\sigma_2}} + \cdots + \beta_J\frac{(X_j - \bar{X}_j)}{\sqrt{\sigma_j}} + \varepsilon \\ \hat{Y}_{PCR} &= \beta_0 + \beta_1X_1 + \beta_2X_2 + \cdots + \beta_JX_J + \varepsilon \end{aligned} \tag{21}$$

3. HASIL DAN PEMBAHASAN

3.1 Pemodelan Regresi dengan Metode *Ordinary Least Square*

Hasil analisis regresi linier berganda dengan menggunakan metode OLS terhadap data IPM, menghasilkan model regresi linier berganda yaitu:

$$\hat{Y} = 4{,}729 \times 10^{-16} + 0{,}1637Z_1 + 0{,}2979Z_2 - 0{,}3130Z_3 + 0{,}3109Z_4 + 0{,}5105Z_5 + 0{,}1277Z_6 + 0{,}0832Z_7$$

3.2. Korelasi Pearson

Uji korelasi digunakan untuk mengetahui hubungan antara 2 variabel yaitu variabel independen dan variabel dependen. Hasil uji korelasi pearson disajikan pada Tabel 2.

Tabel 2. Hasil Uji Korelasi Pearson

Variabel	*p-value*	Keputusan
Z_1	0,5252	Terima H_0
Z_2	3,248e-09	Tolak H_0
Z_3	0,9214	Terima H_0
Z_4	0,003474	Tolak H_0
Z_5	5,684e-13	Tolak H_0
Z_6	0,001775	Tolak H_0
Z_7	0,01645	Tolak H_0

Uji korelasi Pearson dilakukan untuk melihat ada tidaknya hubungan yang signifikan antara masing-masing variabel independen (Z_1 hingga Z_7) terhadap variabel dependen. Berdasarkan hasil uji yang ditunjukkan pada Tabel 2, diperoleh bahwa variabel Z_2, Z_4, Z_5, Z_6, dan Z_7 memiliki nilai p-value lebih kecil dari 0,05, sehingga H0 ditolak. Hal ini menunjukkan bahwa terdapat hubungan yang signifikan antara kelima variabel tersebut dengan variabel dependen. Sebaliknya, untuk variabel Z_1 dan Z_3, nilai p-value masing-masing sebesar 0,5252 dan 0,9214, yang lebih besar dari 0,05. Dengan demikian, H_0 diterima untuk Z_1 dan Z_3, yang berarti tidak terdapat hubungan signifikan antara kedua variabel tersebut terhadap variabel dependen. Hasil ini dapat menjadi pertimbangan dalam pemilihan variabel untuk model regresi selanjutnya.

3.3. Uji Multikolinearitas

Pada penelitian ini uji multikolinearitas digunakan untuk mengetahui apakah terdapat korelasi antar variabel bebas dengan menggunakan nilai VIF. Adapun hasil pengujian multikolinearitas dapat dilihat pada Tabel 3.

Tabel 3. Hasil Uji Multikolinearitas

Variabel	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7
Z_1	Inf	1,15	18,85	1,02	1,02	1,20	2,55
Z_2	1,15	Inf	1,07	1,06	2,13	1,30	1.24
Z_3	18,85	1,07	Inf	1,02	1,00	1,09	1,97
Z_4	1,02	1,06	1,02	Inf	1,03	1,00	1,01
Z_5	1,02	2,13	1,00	1,03	Inf	1,23	1,35
Z_6	1,20	1,30	1,09	1,00	1,23	Inf	1,29
Z_7	2,55	1,24	1,97	1,01	1,35	1,29	Inf

Berdasarkan hasil perhitungan nilai Variance Inflation Factor (VIF) pada Tabel 3, diketahui bahwa sebagian besar nilai VIF antar variabel berada di bawah angka 10, yang mengindikasikan tidak adanya masalah multikolinearitas. Namun, terdapat pengecualian pada pasangan variable Z_1 dan Z_3, di mana nilai VIF mencapai 18,85. Nilai ini melebihi ambang batas umum (VIF > 10), sehingga menunjukkan adanya gejala multikolinearitas yang cukup kuat di antara kedua variabel tersebut. Dengan demikian, H_0 ditolak untuk pasangan Z_1 dan Z_3, yang berarti terjadi korelasi tinggi antar kedua variabel bebas tersebut dan dapat mengganggu kestabilan model regresi. Untuk mengatasi masalah ini, perlu dilakukan penanganan lebih lanjut seperti penghapusan salah satu variabel, transformasi data, atau menggunakan pendekatan regresi alternatif seperti Principal Component Regression (PCR) atau Ridge Regression.

3.4. Pemodelan *Jackknife Ridge Regression*

Pada analisis regresi dengan menggunakan *jackknife ridge regresssion*, data yang digunakan adalah data yang sudah distandarisasi. Langkah pertama yang dilakukan yaitu membuat matriks X^TX dari data yang telah distandarisasi, sehingga akan diperoleh nilai eigen dan vektor eigen dari matriks tersebut. Kemudian menghitung $\hat{\gamma}_{OLS}$ dan $\hat{\sigma}^2_{OLS}$ yang digunakan untuk memperoleh nilai konstanta bias *k* awal.

Tabel 4. Parameter *Ordinary Least Square*

Variabel	$\hat{\gamma}_{OLS}$	$\hat{\sigma}^2_{OLS}$	*k* awal
Z_1	0,2986		0,3144
Z_2	0,6253		0,0717
Z_3	-0,0231		52,7298
Z_4	0,0922	0,0280	3,2982
Z_5	0,0842		3,9505
Z_6	-0,0798		4,9863
Z_7	0,3035		0,3043

Setelah diperoleh nilai konstanta bias *k* awal maka selanjutnya menghitung $\hat{\gamma}_{GRR}$, $\hat{\beta}_{GRR}$ dan $\hat{\sigma}^2_{GRR}$ untuk memperoleh nilai konstanta bias *k* selanjunya. Nilai *k* diperoleh dengan melakukan iterasi HKB dengan persamaan (13). Proses pemilihan nilai *k* dengan iterasi berhenti jika $\left|(\hat{\gamma}_{GRR}{}^{T}\hat{\gamma}_{GRR})^{i}(\hat{\gamma}_{GRR}{}^{T}\hat{\gamma}_{GRR})^{i-1}\right| \leq 0{,}01$. Berdasarkan hasil iterasi, proses iterasi terjadi sebanyak empat kali dan telah memenuhi $\left|(\hat{\gamma}_{GRR}{}^{T}\hat{\gamma}_{GRR})^{i}(\hat{\gamma}_{GRR}{}^{T}\hat{\gamma}_{GRR})^{i-1}\right| \leq 0{,}01$, sehingga nilai *k* pada iterasi keempat yang akan digunakan untuk mengestimasi parameter *jackknife ridge regression*.

Tabel 5. Parameter *Jackknife Ridge Regression*

Variabel	*k*	$\hat{\gamma}_{JRR}$	β_{JRR}
Z_1	0,4174	0,2986	-0,0737
Z_2	0,0948	0,6253	0,3634
Z_3	5521	-2,3539e-08	-0,1440
Z_4	7,9682	0,0843	0,2794
Z_5	25,9472	0,0453	0,4618
Z_6	7607	-6,3166e-07	0,1545
Z_7	34,6409	0,0080	0,1262

Berdasakan Tabel 5, maka diperoleh model untuk *jackknife ridge regression* sebagai berikut.

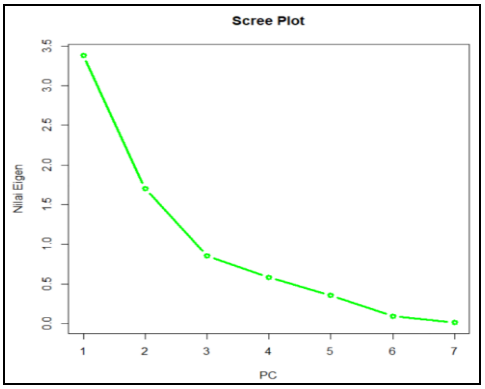
$$\hat{Y} = -0{,}0737Z_1 + 0{,}3634Z_2 - 0{,}1440Z_3 + 0{,}2794Z_4 + 0{,}4618Z_5 + 0{,}1545Z_6 + 0{,}1262Z_7$$

Kemudian mengembalikan persamaan *jackknife ridge regression* kebentuk awal regresi dengan persamaan (16), sehingga model *jackknife ridge regression* menjadi:

$$\hat{Y} = -10{,}0952 - 2{,}4409 \times 10^{-05}X_1 + 0{,}7753X_2 - 5{,}0609 \times 10^{-04}X_3 + 1{,}4348X_4 + 7{,}8172 \times 10^{-04}X_5 + 0{,}3672X_6 + 6{,}1341 \times 10^{-07}X_7$$

3.5. Pemodelan *Principal Component Regression*

Langkah pertama yang dilakukan membuat matriks korelasi (ρ) dan menghitung nilai eigen dan vektor eigen dari matriks tersebut. Nilai eigen yang diperoleh akan digunakan untuk menentukan jumlah komponen utama melalui scree plot. Scree plot yang dihasilkan dapat dilihat bahwa titik plot nilai eigen yang lebih besar dari 1 terdapat pada PC 2 yang artinya terdapat 2 komponen utama yang terbentuk.



Gambar 1. *Scree Plot*

Setelah memperoleh jumlah komponen utama, maka selanjutnya membentuk persamaan komponen utama berdasarkan nilai vektor eigen atau loadings. Hasil perhitungan diperoleh seperti pada Tabel 6.

Tabel 6. Skor Komponen Utama

W_1	W_2
0,4686	-0,3441
0,3739	0,4267
0,4211	-0,4192
-0,0130	0,4538
0,3135	0,5210
0,3592	0,1913
0,4856	-0,0986

.

Berdasarkan Tabel 6 maka persamaan komponen utama yang terbentuk adalah:

$$W_1 = 0{,}4686Z_1 + 0{,}3739Z_2 + 0{,}4211Z_3 - 0{,}0130Z_4 + 0{,}3135Z_5 + 0{,}3592Z_6 + 0{,}4856Z_7$$
$$W_2 = -0{,}3441Z_1 + 0{,}4267Z_2 - 0{,}4192Z_3 + 0{,}4538Z_4 + 0{,}5210Z_5 + 0{,}1913Z_6 - 0{,}0986Z_7$$

Setelah persamaan komponen utama diperoleh dan telah terbentuk variabel baru maka selanjutnya memasukkan variabel baru komponen utama yang terbentuk kedalam data dan akan diregresikan dengan variabel dependen (Y) dengan metode OLS. Maka diperoleh model *principal component regression* yang terbentuk sebagai berikut.

$$\hat{Y} = 1{,}173 \times 10^{-15} + 0{,}2986W_1 + 0{,}6253W_2$$

Kemudian mengembalikan model *principal component regression* kedalam bentuk variabel baku, diperoleh model sebagai berikut.

$$\hat{Y} = 1{,}173 \times 10^{-15} - 0{,}0752Z_1 + 0{,}3784Z_2 - 0{,}1364Z_3 + 0{,}3237Z_4 + 0{,}4194Z_5 + 0{,}2269Z_6 + 0{,}0834Z_7$$

Langkah terakhir dengan mengembalikan model *principal component regression* variabel baku ke variabel asal, diperoleh model sebagai berikut.

$$\hat{Y} = -23{,}4854 - 6{,}5520 \times 10^{-06}X_1 + 0{,}2123X_2 - 1{,}2603 \times 10^{-04}X_3 + 0{,}4370X_4 + 1{,}8667 \times 10^{-04}X_5 + 0{,}1418X_6 + 1{,}0654 \times 10^{-07}X_7$$

3.6. Pemilihan Model Terbaik

Pemilihan model terbaik pada penelitian ini yaitu dengan menggunakan nilai *Adjusted* R^2 dan MSE yang dihasilkan kedua metode tersebut. Kriteria pemilihan model terbaik yaitu model yang memiliki nilai *Adjusted* R^2 terbesar dan MSE terkecil. Nilai *Adjusted* R^2 dan MSE yang dihasilkan disajikan pada Tabel 7.

Tabel 7. Perbandingan Nilai *Adjusted* R^2 dan MSE

Metode	*Adjusted* R^2	MSE
Jackknife Ridge Regression	0,968	0,033
Principal Component Regression	0,966	0,040

Berdasarkan Tabel 7, dapat disimpulkan bahwa model JRR merupakan model terbaik karena memiliki nilai *Adjusted* R^2 lebih besar dan nilai MSE lebih kecil dibandingkan dengan PCR

3.7. Uji Signifikansi

Selanjutnya dilakukan uji F dan uji t untuk mengetahui signifikansi model regresi pada model terbaik. Uji F digunakan untuk mengetahui seberapa signifikan pengaruh variabel independen secara simultan terhadap variabel dependen dengan kriteria pengujian jika $F_{hitung} > F_{(\alpha,j,n-j-1)}$ atau $p - value < \alpha$, maka H_0 ditolak.

Tabel 8. Uji F

F_{hit}	$F_{(0,05;7;26)}$
139,2987	2,39

Berdasarkan hasil uji F yang ditampilkan pada Tabel 8, diperoleh nilai F_{hitung} sebesar 139,2987, sedangkan nilai $F_{(\alpha,j,n-j-1)}$ pada tingkat signifikansi α = 0,05 dengan derajat kebebasan (7; 26) adalah 2,39. Karena $F_{hitung} > F_{(\alpha,j,n-j-1)}$ (139,2987 > 2,39), maka keputusan yang diambil adalah menolak H_0. Artinya, model regresi yang digunakan menunjukkan bahwa variabel independen secara simultan (bersama-sama) memiliki pengaruh yang signifikan terhadap variabel dependen. Dengan kata lain, kombinasi variabel Z1 hingga Z7 secara bersama-sama dapat menjelaskan variasi yang terjadi pada variabel dependen dalam model terbaik yang digunakan.

Setelah uji F menunjukkan signifikansi model secara keseluruhan, maka dilakukan uji t untuk menguji pengaruh masing-masing variabel independen secara parsial terhadap variabel dependen. Kriteria pengujiannya adalah jika nilai absolut $|t_{hitung}| > t_{(\alpha/2,n-j-1)}$ atau $p - value < \alpha$, maka H_0, yang berarti variabel independen tersebut berpengaruh secara signifikan terhadap variabel dependen. Sayangnya, nilai-nilai thitung dan p-value dari masing-masing variabel tidak ditampilkan dalam bagian ini, sehingga interpretasi uji t secara rinci belum dapat dilakukan. Namun, secara umum, jika ada variabel dengan nilai thitung signifikan atau p-value kurang dari 0,05, maka variabel tersebut dapat dinyatakan berpengaruh secara parsial terhadap variabel dependen, dan dapat dipertahankan dalam model regresi akhir.

Dengan demikian, dari uji F dapat disimpulkan bahwa model secara keseluruhan signifikan, dan uji t dapat digunakan untuk memperkuat kesimpulan terkait kontribusi masing-masing variabel independen secara individual.

Tabel 9. Uji t

Variabel	$\|t_{hit}\|$	$t_{(0,025;26)}$	Keputusan
X_1	4,6532		Tolak H_0
X_2	16,277		Tolak H_0
X_3	447297		Tolak H_0
X_4	8,0224	2,056	Tolak H_0
X_5	17,9568		Tolak H_0
X_6	196059,5		Tolak H_0
X_7	19,5797		Tolak H_0

Berdasarkan hasil uji t dapat disimpulkan bahwa tolak H_0, dengan artian semua variabel berpengaruh signifikan secara parsial terhadap variabel dependen karena $|t_{hitung}| > t_{(\alpha/2, n-j-1)}$.

4. KESIMPULAN DAN SARAN/REKOMENDASI

4.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa model terbaik untuk menangani masalah multikolinearitas dalam pemodelan Indeks Pembangunan Manusia (IPM) di Indonesia adalah model dengan metode Jackknife Ridge Regression (JRR). Metode ini terbukti lebih unggul dibandingkan metode Principal Component Regression (PCR), ditunjukkan oleh nilai Adjusted R^2 tertinggi sebesar 96,7% dan nilai Mean Square Error (MSE) terendah sebesar 0,033. Artinya, model JRR mampu menjelaskan hampir seluruh variasi IPM dengan kesalahan prediksi yang sangat kecil.

Selain itu, dari model terbaik tersebut diperoleh bahwa seluruh variabel independen yang digunakan, yaitu jumlah penduduk, angka harapan hidup, jumlah penduduk miskin, harapan lama sekolah, pengeluaran per kapita, tingkat pengangguran terbuka, dan produk domestik regional bruto (PDRB), berpengaruh signifikan secara parsial maupun simultan terhadap IPM. Hal ini menunjukkan bahwa faktor-faktor ekonomi, pendidikan, dan demografi memainkan peran penting dalam menentukan kualitas pembangunan manusia di Indonesia.

4.2 Saran/Rekomendasi

Sebagai tindak lanjut dari penelitian ini, disarankan agar penelitian selanjutnya mengembangkan model dengan membandingkan metode Jackknife Ridge Regression dengan metode regresi lainnya yang juga mampu mengatasi multikolinearitas, seperti Lasso Regression, Elastic Net, atau Partial Least Squares Regression (PLSR).

Selain itu, peneliti selanjutnya juga disarankan untuk menggunakan data dengan cakupan waktu yang lebih panjang (time series) atau lebih luas (cross-section yang lebih banyak), serta menambahkan variabel lain yang relevan, seperti tingkat melek huruf, akses layanan kesehatan, atau infrastruktur daerah, untuk memperoleh gambaran yang lebih komprehensif dalam pemodelan IPM di Indonesia. Dengan demikian, hasil penelitian dapat lebih akurat dan aplikatif sebagai dasar perumusan kebijakan pembangunan yang berkelanjutan.

REFERENSI

Amelia, S., & Putra, A. A. (2023). Regresi Komponen Utama dalam Mengatasi Multikolinearitas pada Faktor yang Mempengaruhi Pendapatan Asli Daerah di Sumatera Barat. *Jurnal Pendidikan Tambusai, 7*, 10906-10914.

Anggraini, N., Kusnandar, D., & Debataraja, N. N. (2019). Metode Generalized Ridge Regression Dalam Mengatasi Multikolinearitas. *Bimaster: Buletin Ilmiah Matematika, Statistika Dan Terapannya*, *8*(4).

Arrasyid, A. H., Ispriyanti, D., & Hoyyi, A. (2021). Metode Modified Jackknife Ridge Regression Dalam Penanganan Multikolinearitas (Studi Kasus: Indeks Pembangunan Manusia Di Jawa Tengah). *JURNAL GAUSSIAN, 10*, 104-113.

Arum, P. R., & Fitri, Y. (2023). Analisis Faktor-Faktor Yang Mempengaruhi Indeks Pembangunan Manusia Berdasarkan Kabupaten/Kota Di Jawa Tengah. *J Statistika: Jurnal Ilmiah Teori Dan Aplikasi Statistika*, *16*(1), 424-433.

Auqino, S., Maiyastri, & Diana, R. (2019). Perbandingan Metode Kuadrat Terkecil Dan Metode Bayes pada Model Regresi Linier Berganda yang Mengandung Multikolinearitas. *Jurnal Matematika UNAND, 8*, 307-312.

Bau, M. K., Setyawan, Y., & Jatipaningrum, M. T. (2023). Perbandingan Metode Algoritma K-Means Dan K-Medoids Pada Pengelompokan Kabupaten/Kota Di Provinsi Nusa Tenggara Timur Berdasarkan

Dimensi Indeks Pembangunan Manusia Tahun 2020. *Jurnal Statistika Industri dan Komputasi, 8*, 48-57.

Enzellina, G., & Suhaedi, D. (2022). Penggunaan Metode Principal Component Analysis Dalam Menentukan Faktor Dominan. *Jurnal Riset Matematika*, 101-110.

Fatihah, A. I. (2022). *Penerapan Metode Jackknife Ridge Regression Dalam Penanganan Multikolinearitas (Studi Kasus: Tingkat Kemiskinan di Kabupaten Grobogan Tahun 2011-2021)*. Semarang: Universitas Islam Negeri Walisongo.

Fauzan, E. A., & Soehardjoepri, S. (2023). Perbandingan Metode Regresi Komponen Utama dan Regresi Ridge pada Analisis Pengaruh Rasio Keuangan Profitabilitas terhadap Harga Saham Perusahaan Perbankan. *Jurnal Sains dan Seni ITS*, *11*(6), D367-D374.

Gujarati, D. N., & C.Porter, D. (2009). *Basic Econometrics.* McGraw-hill: Douglas Reiner.

Khoiri, N. (2021). *Buku Statistika Konseptual dan Aplikasi Perspektif Manjemen.* Southeast Asian Publishing.

Maulida, R. (2022). *Perbandingan Principal Component Regression Dan Regresi Ridge Pada Analisis Faktor-Faktor Indeks Pembangunan Manusia.* Malang: Universitas Islam Negeri Maulana Malik Ibrahim.

Ningrum, J. W., Khairunnisa, A. H., & Huda, N. (2020). Pengaruh Kemiskinan, Tingkat Pengangguran, Pertumbuhan Ekonomi dan Pengeluaran Pemerintah Terhadap Indeks Pembangunan Manusia (IPM) di Indonesia Tahun 2014-2018 dalam Perspektif Islam. *Jurnal Ilmiah Ekonomi Islam, 6*, 212-222.

Pendi, P. (2021). Analisis Regresi Dengan Metode Komponen Utama Dalam Mengatasi Masalah Multikolinearitas. *Bimaster Buletin Ilmiah Matematika Statistika dan Terapannya*, 131-138.

Rahmadeni, & Yunita, V. (2019). Pemodelan Indeks Pembangunan Manusia di Provinsi Riau dengan Menggunakan Regresi Logistik Ordinal. *Jurnal Sains Matematika dan Statistika, 5*, 120-126.

Roflin, E., & Zulvia, F. E. (2021). *Kupas Tuntas Analisis Korelasi* . Penerbit NEM.

Rusli, Y. N., & Setyawan, Y. (2022). Analisis Faktor-Faktor Kesenjangan Sosial Daerah Istimewa Yogyakarta Menggunakan Metode Principal Component Analysis (Pca) Dan Confirmatory Factor Analysis (Cfa). *Jurnal Statistika Industri dan Komputasi, 7*, 41-53.

Sangkereng, W., Engka, D. S., & Sumual, J. I. (2019). Faktor-Faktor Yang Mempengaruhi Indeks Pembangunan Manusia Di Provinsi Sulawesi Utara. *Jurnal Berkala Ilmiah Efisiensi, 19*, 60-71.

Siburian, J. N., Rahmawati, R., & Hoyyi, A. (2019). Regresi komponen utama robusts-estimator untuk analisis pengaruh jumlah pengangguran di Jawa Tengah. *Jurnal Gaussian, 8*, 439-450.

Si'Iang, I. L., Hasid, Z., & Priyagus. (2019). Analisis faktor-faktor yang berpengaruh terhadap indeks pembangunan manusia. *Jurnal Manajemen, 11*, 159-169.

Silitonga, Y. C., Kamid, & Multahadah, C. (2021). Perbandingan Metode Stepwise Dan Principal Componen Analysis (Pca) pada Kasus Faktor-Faktor yang Mempengaruhi Pendapatan Asli Daerah (PAD) Di Provinsi Jambi. *Jurnal Gamma-Pi, 3*, 12-20.

Syarifuddin, S., & Saudi, I. A. (2022). *Metode Riset Praktis Regresi Berganda Menggunakan SPSS.*

Taqiyyuddin, T. A., & Irfan, M. (2022). Faktor Penyebab Kemiskinan di Provinsi Jawa Barat Menggunakan Spatial Autoregressive Quantile Regression. *Jurnal Sains Matematika dan Statistika, 8*, 59-69.