E-ISSN: 3025-0994
Halaman 1323-1336
Volume 4 Nomor 4 Tahun 2026

Analisis Faktor-Faktor yang Berhubungan dengan Status Kemiskinan Tinggi pada Kabupaten/Kota di Provinsi Jawa Barat Periode 2020-2023 Menggunakan Regresi Logistik Biner

Angel Patricia Insani Gultom[1], Brian Hardiwan Gea[2], Feyza Zaki Nandana[3], Finny Aprilia Yulanda[4], Indri Theresia Manik[5]

[1,2,3,4,5] Program Studi Statistika, Universitas Negeri Medan
e-mail: indrimanik52@gmail.com

Abstrak

Kemiskinan menjadi salah satu persoalan pembangunan yang bersifat kompleks karena berkaitan dengan beragam aspek ekonomi dan demografi. Penelitian ini dilakukan berdasarkan adanya perbedaan kondisi kemiskinan antarwilayah kabupaten dan kota di Provinsi Jawa Barat pada periode 2020–2023. Tujuan penelitian adalah mengidentifikasi faktor-faktor yang berkaitan dengan status kemiskinan tinggi pada kabupaten dan kota di Provinsi Jawa Barat. Penelitian ini menerapkan pendekatan kuantitatif dengan memanfaatkan data sekunder yang bersumber dari Badan Pusat Statistik Provinsi Jawa Barat. Data penelitian melibatkan 27 kabupaten dan kota selama empat tahun pengamatan dengan jumlah keseluruhan 108 observasi. Faktor yang dianalisis mencakup tingkat pengangguran terbuka, tingkat partisipasi angkatan kerja, indeks pembangunan manusia, serta jumlah penduduk. Proses analisis menggunakan regresi logistik biner yang meliputi pemeriksaan data, analisis deskriptif, pengujian multikolinearitas, pengujian signifikansi, interpretasi peluang, serta penilaian kinerja model. Hasil analisis menunjukkan bahwa model penelitian secara keseluruhan memiliki hubungan yang signifikan dengan status kemiskinan tinggi. Pada pengujian secara parsial, indeks pembangunan manusia merupakan satu-satunya variabel yang memiliki hubungan signifikan. Semakin tinggi indeks pembangunan manusia, semakin rendah peluang suatu kabupaten atau kota termasuk dalam kategori kemiskinan tinggi. Sebaliknya, tingkat pengangguran terbuka, tingkat partisipasi angkatan kerja, dan jumlah penduduk tidak menunjukkan hubungan yang signifikan secara parsial. Model menghasilkan ketepatan klasifikasi sebesar 81,48 persen dan kemampuan diskriminasi yang baik. Oleh karena itu, peningkatan pembangunan manusia menjadi salah satu aspek penting dalam menekan status kemiskinan tinggi di Provinsi Jawa Barat.

Kata Kunci: Kemiskinan, Jawa Barat, Indeks Pembangunan Manusia, Regresi Logistik

Abstract

Poverty was one of the complex development problems because it was associated with various economic and demographic aspects. This study was conducted based on differences in

poverty conditions among regencies and cities in West Java Province during the 2020–2023 period. The study aimed to identify factors associated with high poverty status among regencies and cities in West Java Province. A quantitative approach was applied using secondary data obtained from the Statistics Indonesia of West Java Province. The data covered 27 regencies and cities over four years of observation, resulting in a total of 108 observations. The factors analyzed included the open unemployment rate, labor force participation rate, human development index, and population size. The analysis was conducted using binary logistic regression, involving data examination, descriptive analysis, multicollinearity testing, significance testing, odds interpretation, and model performance evaluation. The results show that the overall model has a significant relationship with high poverty status. In the partial analysis, the human development index is the only variable that shows a significant relationship. A higher human development index is associated with a lower probability of a regency or city being classified as having high poverty status. Meanwhile, the open unemployment rate, labor force participation rate, and population size do not show significant partial relationships. The model achieves a classification accuracy of 81.48 percent and demonstrates good discriminatory ability. Therefore, improving human development is an important aspect of reducing high poverty status in West Java Province

Keywords : Poverty, West Java, Human Development Index, Logistic Regression

LATAR BELAKANG

Kemiskinan merupakan masalah pembangunan yang kompleks karena menggambarkan ketidakmampuan individu atau rumah tangga secara ekonomi untuk memenuhi kebutuhan dasar minimal, baik kebutuhan makanan maupun nonmakanan. Badan Pusat Statistik (BPS) menggunakan pendekatan kebutuhan dasar untuk mengukur indikator ini, di mana penduduk dengan rata-rata pengeluaran per kapita di bawah garis kemiskinan diklasifikasikan sebagai penduduk miskin. Tingkat kemiskinan pada suatu wilayah tidak hanya dipengaruhi oleh faktor pendapatan semata, melainkan terkait erat dengan berbagai kondisi makroekonomi dan demografi seperti ketersediaan lapangan kerja, tingkat pendidikan, kesehatan, partisipasi angkatan kerja, dan jumlah penduduk.

Provinsi Jawa Barat, yang terdiri atas 27 kabupaten/kota, memiliki kondisi wilayah yang sangat bervariasi. Berdasarkan data BPS Provinsi Jawa Barat periode 2020–2023, persentase penduduk miskin antarkabupaten/kota menunjukkan rentang yang cukup lebar, yakni berkisar antara 2,38 persen hingga 13,13 persen. Untuk mempermudah identifikasi wilayah yang masuk dalam kriteria rawan atau prioritas penanganan, data persentase penduduk miskin yang bersifat kontinu tersebut dapat dikelompokkan menjadi variabel kategorik biner. Dengan menetapkan batas 10 persen, suatu wilayah diklasifikasikan ke dalam status kemiskinan tinggi atau tidak tinggi. Terdapat fluktuasi jumlah kabupaten/kota dengan status kemiskinan tinggi selama periode observasi, yaitu sebanyak 9 wilayah pada 2020, 12 wilayah pada 2021, 10 wilayah pada 2022, dan 8 wilayah pada 2023. Secara agregat, 39 dari 108 pengamatan (36,11 persen) tergolong berstatus kemiskinan tinggi.

Beberapa indikator diduga memiliki hubungan signifikan terhadap tingginya status kemiskinan di wilayah tersebut. Tingkat Pengangguran Terbuka (TPT) yang tinggi menunjukkan banyaknya angkatan kerja yang tidak memiliki sumber pendapatan, sehingga berdampak pada penurunan daya beli rumah tangga. Tingkat Partisipasi Angkatan Kerja (TPAK) yang menggambarkan proporsi penduduk usia kerja aktif memiliki dampak yang sangat bergantung pada ketersediaan lapangan kerja produktif. Selain itu, Indeks Pembangunan Manusia (IPM) mencerminkan kualitas pendidikan dan kesehatan yang berkontribusi langsung pada kapabilitas dan produktivitas tenaga kerja. Pertumbuhan jumlah penduduk yang besar tanpa diimbangi oleh kapasitas ekonomi dan perluasan lapangan kerja juga berisiko menambah tekanan pada sumber daya serta memperbesar kelompok populasi rentan.

Mengingat variabel respon terkait status kemiskinan dalam analisis ini hanya memiliki dua kategori (tinggi dan tidak tinggi), metode statistika yang paling sesuai untuk digunakan adalah regresi logistik biner. Metode ini merupakan bagian dari kerangka *Generalized Linear Model* (GLM) yang digunakan untuk menganalisis hubungan antara variabel respon berskala biner dengan satu atau lebih variabel prediktor.

Berdasarkan identifikasi masalah tersebut, penelitian ini bertujuan untuk menganalisis faktor-faktor yang berhubungan dengan status kemiskinan tinggi pada kabupaten/kota di Provinsi Jawa Barat selama periode 2020–2023 menggunakan pendekatan regresi logistik biner. Analisis dilakukan terhadap 108 observasi menggunakan data sekunder dari BPS Provinsi Jawa Barat, dengan melibatkan variabel prediktor berupa TPT bulan Agustus, TPAK bulan Agustus, IPM, dan jumlah penduduk.

METODE

Jenis dan Pendekatan Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan deskriptif dan analitik. Penelitian bertujuan menganalisis faktor-faktor yang berhubungan dengan status kemiskinan tinggi pada kabupaten/kota di Provinsi Jawa Barat menggunakan regresi logistik biner. Metode ini digunakan karena variabel dependen memiliki dua kategori, yaitu status kemiskinan tinggi dan tidak tinggi. Analisis dilakukan menggunakan RStudio dengan taraf signifikansi 5%.

Sumber dan Data Penelitian

Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Jawa Barat. Data mencakup 27 kabupaten/kota selama periode 2020–2023, sehingga diperoleh 108 observasi. Variabel yang digunakan meliputi persentase penduduk miskin, TPT Agustus, TPAK Agustus, IPM, dan jumlah penduduk. Data disusun berdasarkan kabupaten/kota dan tahun pengamatan.

Variabel Penelitian

Variabel dependen adalah status kemiskinan tinggi, yang dikategorikan menjadi 1 = tinggi dan 0 = tidak tinggi berdasarkan persentase penduduk miskin. Batas kategorisasi yang digunakan adalah 10%, yaitu persentase penduduk miskin ≥ 10% dikategorikan tinggi dan < 10% dikategorikan tidak tinggi. Variabel independen terdiri atas TPT Agustus (X_1), TPAK Agustus (X_2), IPM (X_3), dan logaritma natural jumlah penduduk (X_4). Persentase penduduk

miskin tidak dimasukkan sebagai prediktor karena digunakan untuk membentuk variabel dependen.

Teknik Analisis Data

Analisis dilakukan melalui beberapa tahap, yaitu: (1) pemeriksaan dan pembersihan data, meliputi pengecekan struktur data, missing value, dan duplikasi; (2) pembentukan variabel status kemiskinan; (3) analisis statistik deskriptif; (4) pemeriksaan multikolinearitas menggunakan Variance Inflation Factor (VIF); (5) pembentukan model regresi logistik biner; (6) pengujian signifikansi model secara simultan menggunakan Likelihood Ratio Test (LRT) dan secara parsial menggunakan Uji Wald; (7) interpretasi koefisien melalui Odds Ratio (OR); serta (8) evaluasi model menggunakan Hosmer–Lemeshow, pseudo-R^2, AIC/BIC, confusion matrix, akurasi, sensitivitas, spesifisitas, F1-score, dan AUC.

Model regresi logistik biner yang digunakan adalah:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \quad (1)$$

dengan p merupakan probabilitas suatu kabupaten/kota memiliki status kemiskinan tinggi, X_1= TPT Agustus, X_2= TPAK Agustus, X_3= IPM, dan X_4= ln jumlah penduduk.

HASIL DAN PEMBAHASAN

Deskripsi Data

Penelitian ini memakai 108 observasi yang berasal dari 27 kabupaten/kota di Jawa Barat selama periode 2020–2023, sehingga setiap wilayah diamati sebanyak empat kali (27 × 4 = 108). Dengan demikian struktur datanya seimbang. Pemeriksaan awal tidak menemukan nilai yang hilang pada variabel mana pun, dan tidak ada baris yang tercatat ganda.

Terdapat lima variabel yang terlibat, yaitu status kemiskinan tinggi (Y) sebagai variabel respon biner serta TPT, TPAK, IPM, dan ln(jumlah penduduk) sebagai prediktor. Penjelasan masing-masing variabel disajikan pada Tabel 4.1.

Tabel 1. Deskripsi Variabel Penelitian

Variabel	Simbol	Keterangan	Skala/Satuan
Status kemiskinan tinggi	Y	1 = Tinggi (persentase penduduk miskin ≥ 10%), 0 = Tidak Tinggi (< 10%)	Nominal (biner)
TPT	X_1	Tingkat Pengangguran Terbuka bulan Agustus	Persen (%)
TPAK	X_2	Tingkat Partisipasi Angkatan Kerja bulan Agustus	Persen (%)
IPM	X_3	Indeks Pembangunan Manusia	Poin indeks
ln(penduduk)	X_4	Logaritma natural jumlah penduduk (ribu jiwa)	ln (ribu jiwa)

Y diberi nilai 1 untuk kabupaten/kota dengan persentase penduduk miskin minimal 10%, dan 0 untuk yang berada di bawah 10%. Pembagian ini terlihat jelas pada data: kelompok Tidak Tinggi berada pada kisaran 2,38% sampai 9,98%, sedangkan kelompok Tinggi pada kisaran 10,03% sampai 13,13%. Karena Y dibentuk langsung dari persentase penduduk miskin, variabel tersebut tidak disertakan sebagai prediktor; jika disertakan, model akan mengalami pemisahan sempurna (perfect separation). Adapun jumlah penduduk dalam ribu jiwa diwakili oleh logaritma naturalnya, ln(penduduk), sebab skalanya jauh lebih besar dan lebih menyebar dibandingkan prediktor yang lain.

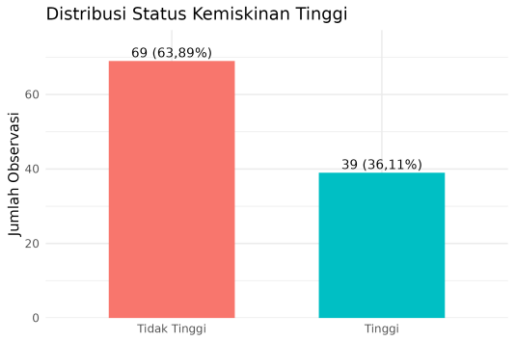
Distribusi Status Kemiskinan Tinggi

Untuk mengetahui proporsi setiap kategori pada variabel respon, disusun distribusi frekuensi yang ditampilkan pada Tabel 2 dan Gambar 1.

Tabel 2. Distribusi Status Kemiskinan Tinggi

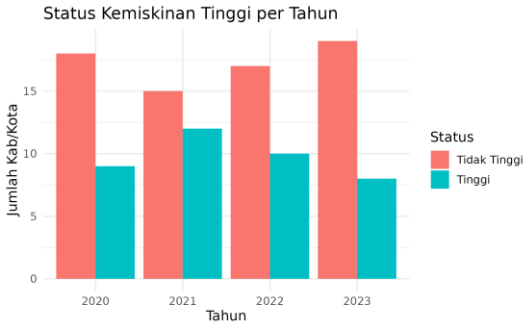
Kategori	Frekuensi	Persentase (%)
Tidak Tinggi (0)	69	63,89
Tinggi (1)	39	36,11
Total	**108**	**100,00**

Dari 108 observasi, 69 di antaranya (63,89%) termasuk kategori Tidak Tinggi dan 39 (36,11%) termasuk kategori Tinggi, atau kira-kira 1,8 berbanding 1. Jumlah pada kedua kategori cukup memadai dan selisihnya tidak terlalu jauh, sehingga regresi logistik biner dapat diterapkan pada data ini.



Gambar 1. Distribusi Status Kemiskinan Tinggi

Bila dirinci menurut tahun, persentase kabupaten/kota berstatus Tinggi adalah 33,33% (9 dari 27) pada 2020, 44,44% (12 dari 27) pada 2021, 37,04% (10 dari 27) pada 2022, dan 29,63% (8 dari 27) pada 2023. Angka tertinggi terjadi pada 2021 dan terendah pada 2023. Pada setiap tahun, jumlah wilayah berstatus Tidak Tinggi tetap lebih banyak, seperti terlihat pada Gambar 4.2.



Gambar 2. Status Kemiskinan Tinggi per Tahun

Statistik Deskriptif Variabel Prediktor

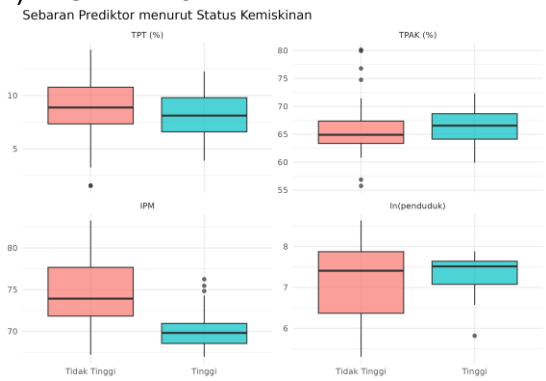
Statistik deskriptif digunakan untuk memberikan gambaran awal tentang pemusatan dan penyebaran setiap variabel prediktor pada 108 observasi, sebagaimana tercantum pada Tabel 4.3.

Tabel 3. Statistik Deskriptif Variabel Prediktor (N = 108)

Variabel	Min	Q1	Median	Mean	Q3	Maks	SD
TPT (%)	1,520	7,085	8,780	8,592	10,557	14,290	2,538
TPAK (%)	55,740	63,722	65,370	65,877	68,228	80,150	3,836
IPM	66,960	69,697	72,500	73,227	75,677	83,290	4,374
ln(penduduk)	5,303	6,941	7,507	7,249	7,829	8,635	0,795

TPT memiliki OR 0,980 dan TPAK OR 0,905; keduanya tidak signifikan karena interval kepercayaan mencakup 1.

Jika dibandingkan menurut status, rata-rata IPM kelompok Tidak Tinggi (74,92) lebih tinggi daripada kelompok Tinggi (70,24). Pada variabel lain selisihnya kecil, yaitu TPT 8,784% berbanding 8,252%, TPAK 65,61% berbanding 66,34%, dan ln(penduduk) 7,199 berbanding 7,338 (masing-masing untuk kelompok Tidak Tinggi dan Tinggi). Pola ini juga tampak pada diagram kotak garis (boxplot) di Gambar 3.



Gambar 3. Sebaran Prediktor menurut Status Kemiskinan

Pada Gambar 3, kotak IPM untuk kelompok Tinggi berada jauh di bawah kelompok Tidak Tinggi, sementara kotak TPT, TPAK, dan ln(penduduk) pada kedua kelompok banyak yang saling tumpang tindih. Beberapa titik pencilan terlihat pada TPT, TPAK, IPM, dan ln(penduduk).

Pemeriksaan Multikolinearitas

Sebelum model ditafsirkan, perlu dipastikan bahwa antarprediktor tidak saling berkorelasi terlalu kuat, karena kondisi tersebut dapat membuat estimasi koefisien menjadi tidak stabil. Pemeriksaan memakai Variance Inflation Factor (VIF) dan Tolerance yang dihitung dari model regresi logistik. Multikolinearitas yang serius ditandai dengan VIF di atas 10 atau Tolerance di bawah 0,1. Hasilnya tercantum pada Tabel 4.

Tabel 4. Hasil Pemeriksaan Multikolinearitas

Variabel	VIF	Tolerance	Keterangan
TPT	2,247	0,4449	Tidak ada multikolinearitas
TPAK	2,291	0,4366	Tidak ada multikolinearitas
IPM	1,521	0,6573	Tidak ada multikolinearitas
ln(penduduk)	1,473	0,6788	Tidak ada multikolinearitas

Nilai VIF keempat prediktor berkisar antara 1,473 dan 2,291, jauh di bawah 10 dan bahkan di bawah 5, sedangkan nilai Tolerance-nya semuanya di atas 0,1. Korelasi terkuat dijumpai antara TPT dan TPAK (−0,661), sementara pasangan lainnya memiliki korelasi absolut paling besar 0,275. Jadi, tidak ada indikasi multikolinearitas yang serius dan keempat prediktor layak dimasukkan bersama-sama ke dalam model.

Pemodelan Regresi Logistik Biner

Model dibentuk dengan regresi logistik biner berfungsi hubung logit. Y (1 = Tinggi, 0 = Tidak Tinggi) menjadi variabel respon, sedangkan TPT, TPAK, IPM, dan ln(penduduk) menjadi prediktornya. Parameter diestimasi dengan Maximum Likelihood Estimation (MLE) melalui algoritma Fisher scoring yang konvergen setelah 6 iterasi. Hasil estimasinya tersaji pada Tabel 4.5.

Tabel 5. Hasil Estimasi Parameter Model Regresi Logistik Biner

Parameter	Koefisien (β)	Std. Error
Konstanta	46,92129	13,28322
TPT	−0,01975	0,13632
TPAK	−0,09945	0,08631
IPM	−0,52549	0,12092
ln(penduduk)	−0,40235	0,36863

Berdasarkan Tabel 5, model yang diperoleh adalah sebagai berikut.

$$ln\big(\pi/(1-\pi)\big) = 46{,}92129 - 0{,}01975TPT - 0{,}09945TPAK - 0{,}52549IPM - 0{,}40235\ ln(penduduk) \quad (2)$$

Dengan π menyatakan peluang suatu kabupaten/kota berstatus kemiskinan tinggi, yang dihitung dari $\pi = exp(\eta)/\big(1 + exp(\eta)\big)$, dan η adalah nilai logit pada persamaan di atas. Semua koefisien prediktor bernilai negatif, artinya kenaikan TPT, TPAK, IPM, maupun ln(penduduk) diikuti penurunan log odds suatu wilayah berstatus kemiskinan tinggi, dengan variabel lain dianggap tetap. Koefisien IPM paling menonjol jika dibandingkan dengan standard error-nya.

Uji Signifikansi Model

Uji signifikansi model dilakukan untuk mengetahui apakah variabel TPT, TPAK, IPM, dan ln(penduduk) secara bersama-sama berhubungan dengan status kemiskinan tinggi. Pengujian dilakukan menggunakan Likelihood Ratio Test (LRT) atau uji G dengan taraf signifikansi 5%.

Hipotesis yang digunakan adalah H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, sedangkan H1: paling sedikit terdapat satu β_j yang tidak sama dengan nol.

Tabel 6. Hasil Uji Signifikansi Model (Likelihood Ratio Test)

Statistik	Nilai
Null Deviance	141,28
Residual Deviance	100,69
G-hitung	40,589
Derajat bebas	4
χ^2 tabel (α = 0,05)	9,488
p-value	$3{,}269 \times 10^{-8}$

Hasil pengujian menunjukkan bahwa nilai G-hitung sebesar 40,589 lebih besar daripada χ^2 tabel sebesar 9,488. Nilai p-value sebesar $3{,}269 \times 10^{-8}$ juga lebih kecil dari 0,05. Oleh karena itu, H0 ditolak. Hasil ini menunjukkan bahwa TPT, TPAK, IPM, dan ln(penduduk) secara bersama-sama berhubungan secara signifikan dengan status kemiskinan tinggi. Dengan demikian, model yang digunakan memberikan informasi yang lebih baik dibandingkan model tanpa variabel prediktor.

Uji Signifikansi Parsial

Uji signifikansi parsial dilakukan menggunakan uji Wald untuk melihat hubungan masing-masing variabel prediktor dengan status kemiskinan tinggi setelah variabel lain dalam model diperhitungkan.

Hipotesis untuk masing-masing variabel adalah H_0: $\beta_j = 0$ dan H_1: $\beta_j \neq 0$. Hasil uji parsial disajikan pada Tabel berikut.

Tabel 7. Hasil Uji Signifikansi Parsial (Wald)

Variabel	Koefisien	Std. Error	Wald z	p-value	Keputusan
TPT	-0,01975	0,13632	-0,1449	0,8848	Tidak signifikan
TPAK	-0,09945	0,08631	-1,1522	0,2492	Tidak signifikan
IPM	-0,52549	0,12092	-4,3459	0,0000139	Signifikan
ln(penduduk)	-0,40235	0,36863	-1,0915	0,2751	Tidak signifikan

Dari hasil tersebut, hanya IPM yang memiliki p-value kurang dari 0,05. Artinya, IPM menunjukkan hubungan yang signifikan dengan status kemiskinan tinggi. Sementara itu, TPT, TPAK, dan ln(penduduk) tidak menunjukkan hubungan yang signifikan secara parsial pada taraf 5%. Hasil ini juga sejalan dengan pengujian likelihood ratio per variabel, yang menunjukkan bahwa IPM merupakan satu-satunya prediktor yang signifikan.

Interpretasi Odds Ratio

Odds ratio digunakan untuk mempermudah interpretasi koefisien pada regresi logistik. Nilai OR diperoleh dari exp(β). Nilai OR di bawah satu menunjukkan arah hubungan negatif, sedangkan nilai di atas satu menunjukkan arah hubungan positif.

Tabel 8. Hasil Odds Ratio, Interval Kepercayaan 95%, dan p-value

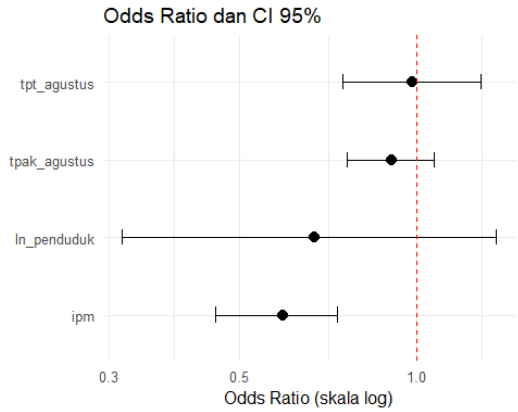
Variabel	OR	CI 95%	p-value
TPT	0,980	0,748–1,284	0,8848
TPAK	0,905	0,760–1,070	0,2492
IPM	0,591	0,455–0,731	0,0000139
ln(penduduk)	0,669	0,316–1,364	0,2751

TPT memiliki OR 0,980 dan TPAK OR 0,905; keduanya tidak signifikan karena interval kepercayaan mencakup 1.

TPAK memiliki OR sebesar 0,905. Artinya, kenaikan satu satuan TPAK berkaitan dengan penurunan odds status kemiskinan tinggi sekitar 9,47%, dengan variabel lain konstan. Hubungan tersebut juga tidak signifikan karena p-value sebesar 0,2492.

IPM memiliki OR 0,591 (95% CI: 0,455–0,731). Setiap kenaikan satu poin IPM berkaitan dengan penurunan odds status kemiskinan tinggi sekitar 40,87%.

ln(penduduk) memiliki OR 0,669 dan tidak signifikan karena interval kepercayaannya mencakup 1.



Gambar 4. Odds Ratio dan Interval Kepercayaan 95% Prediktor

Secara visual, interval kepercayaan IPM berada seluruhnya di bawah garis OR = 1, sedangkan interval prediktor lainnya masih melintasi garis tersebut. Pola ini memperjelas bahwa IPM merupakan prediktor yang memiliki bukti hubungan paling konsisten dalam model, sementara prediktor lainnya belum menunjukkan pola yang meyakinkan secara statistik.

Evaluasi Model

Uji Hosmer–Lemeshow menghasilkan χ^2 = 12,968; df = 8; p = 0,113, sehingga tidak terdapat bukti ketidaksesuaian model.

Uji Hosmer–Lemeshow menghasilkan χ^2 = 12,968; df = 8; p = 0,113, sehingga tidak terdapat bukti ketidaksesuaian model.

Pseudo R^2, AIC, dan BIC

Tabel 9. Nilai Pseudo R^2, AIC, dan BIC

Ukuran	Nilai
McFadden R^2	0,2873
Cox–Snell R^2	0,3133
Nagelkerke R^2	0,4293
AIC	110,7
BIC	124,1

Nilai Nagelkerke R^2 sebesar 0,4293 menunjukkan kemampuan penjelasan model berdasarkan ukuran pseudo R^2 tersebut. Nilai AIC dan BIC dapat digunakan sebagai ukuran pembanding apabila terdapat model alternatif dengan spesifikasi yang berbeda.

Confusion Matrix

Tabel 10. Confusion Matrix Model Regresi Logistik

Aktual / Prediksi	Tidak Tinggi	Tinggi
Tidak Tinggi	60	9
Tinggi	11	28

Tabel 11. Metrik Evaluasi Klasifikasi Model

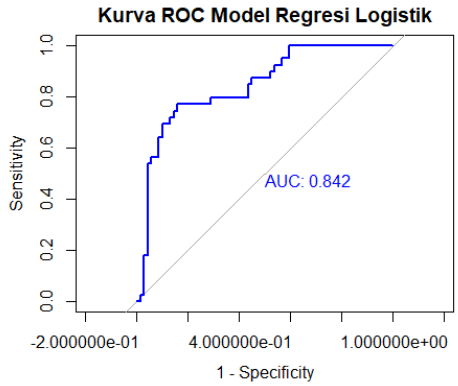
Ukuran	Nilai
Akurasi	81,48%
Sensitivitas	71,79%
Spesifisitas	86,96%
Presisi	75,68%
F1-score	73,68%

Dengan cut-off probabilitas 0,5, model mengklasifikasikan 81,48% observasi dengan benar. Sensitivitas sebesar 71,79% menunjukkan kemampuan model dalam mengenali observasi yang berstatus kemiskinan tinggi, sedangkan spesifisitas sebesar 86,96% menunjukkan kemampuan model dalam mengenali observasi yang berstatus tidak tinggi.

ROC-AUC

Nilai AUC sebesar 0,8417 (95% CI: 0,7635–0,9198), menunjukkan kemampuan diskriminasi model yang baik. Cut-off Youden 0,4634 menghasilkan sensitivitas 76,92% dan spesifisitas 84,06%.

Sebagai analisis pembanding, cut-off optimal berdasarkan metode Youden diperoleh sebesar 0,4634, dengan sensitivitas 76,92% dan spesifisitas 84,06%. Namun, penelitian ini tetap menggunakan cut-off 0,5 sebagai batas klasifikasi utama.



1. **Gambar 5. Kurva ROC Model Regresi Logistik**

Secara visual, kurva ROC berada cukup jauh di atas garis diagonal sehingga model menunjukkan kemampuan yang baik dalam membedakan kedua kategori status kemiskinan. Bentuk kurva yang mengarah ke sudut kiri atas juga memperlihatkan bahwa model memiliki daya diskriminasi yang lebih baik daripada klasifikasi secara acak.

PEMBAHASAN

Berdasarkan hasil analisis, model regresi logistik biner yang digunakan secara keseluruhan menunjukkan hubungan yang signifikan dengan status kemiskinan tinggi. Hal ini terlihat dari hasil uji G sebesar 40,589 dengan p-value $3{,}269 \times 10^{-8}$. Artinya, variabel TPT, TPAK, IPM, dan ln(penduduk) secara bersama-sama memberikan informasi yang berkaitan dengan status kemiskinan tinggi.

Jika dilihat satu per satu, IPM menjadi satu-satunya variabel yang menunjukkan hubungan signifikan. Koefisien IPM bernilai negatif sebesar -0,52549 dengan OR 0,591. Hasil ini menunjukkan bahwa kabupaten/kota dengan IPM yang lebih tinggi cenderung memiliki odds untuk berada pada status kemiskinan tinggi yang lebih rendah setelah memperhitungkan TPT, TPAK, dan ln(penduduk).

Sementara itu, TPT, TPAK, dan ln(penduduk) tidak menunjukkan hubungan yang signifikan pada taraf 5%. Walaupun ketiga variabel tersebut memiliki arah koefisien negatif, arah tersebut belum cukup untuk dijadikan dasar dalam menyatakan adanya hubungan yang signifikan secara statistik.

Dari evaluasi model, uji Hosmer–Lemeshow menghasilkan p-value 0,113 sehingga tidak terdapat bukti ketidaksesuaian model yang signifikan. Model juga menghasilkan Nagelkerke R^2 sebesar 0,4293. Pada cut-off 0,5, akurasi model mencapai 81,48% dan AUC sebesar 0,8417. Hasil tersebut menunjukkan bahwa model memiliki kemampuan klasifikasi yang cukup baik pada data yang dianalisis.

Namun, hasil penelitian perlu dipahami dengan memperhatikan struktur data. Data terdiri atas 27 kabupaten/kota yang diamati selama empat tahun, sehingga pengamatan dari kabupaten/kota yang sama pada tahun berbeda kemungkinan tidak sepenuhnya independen. Oleh karena itu, hasil regresi ini lebih tepat dipandang sebagai analisis hubungan pada data yang digunakan, bukan sebagai bukti hubungan sebab-akibat. Untuk penelitian selanjutnya, struktur data berulang dapat dianalisis menggunakan pendekatan yang secara khusus mempertimbangkan efek wilayah dan waktu, seperti regresi logistik data panel, GEE, atau model efek campuran.

SIMPULAN

Hasil penelitian memperlihatkan bahwa model regresi logistik biner secara simultan memiliki hubungan yang signifikan dengan status kemiskinan tinggi di Provinsi Jawa Barat. Secara parsial, indeks pembangunan manusia merupakan satu-satunya variabel yang menunjukkan hubungan signifikan, sementara tingkat pengangguran terbuka, tingkat partisipasi angkatan kerja, dan jumlah penduduk tidak menunjukkan hubungan yang signifikan. Model mampu mengklasifikasikan data dengan tingkat ketepatan sebesar 81,48 persen. Pemerintah Provinsi Jawa Barat serta pemerintah kabupaten/kota perlu memperkuat

pendidikan, kesehatan, dan kapasitas masyarakat sebagai upaya mengurangi kemiskinan. Penelitian berikutnya dapat mempertimbangkan aspek wilayah dan waktu dalam analisis.

DAFTAR PUSTAKA

Adnan, M., & Firmansyah, F. (2022). Analisis Pengaruh Pdrb Per Kapita Dan Tingkat Pendidikan Terhadap Tingkat Pengangguran Terbuka Di Indonesia. Jurnal Ekonomi Dan Pembangunan Regional, 13(2), 101-114.

Agresti, A. (2019). *An Introduction To Categorical Data Analysis* (3rd Ed.). John Wiley & Sons. An Introduction To Categorical Data Analysis - Alan Agresti - Google Buku

Agresti, A. (2019). An Introduction To Categorical Data Analysis (3rd Ed.). John Wiley & Sons.

Bhakti, R., & Handayani, S. (2023). Pemodelan Regresi Logistik Biner Untuk Klasifikasi Kabupaten/Kota Berdasarkan Tingkat Pengangguran Di Pulau Sumatra. Jurnal Sains Data Dan Statistik, 4(1), 33-45.

Dewi, N. L., & Utama, M. S. (2021). Pengaruh Indeks Pembangunan Manusia, Jumlah Penduduk, Dan Pdrb Terhadap Tingkat Pengangguran Terbuka. E-Jurnal Ekonomi Pembangunan Universitas Udayana, 10(3), 882-910.

Ehrenberg, R. G., & Smith, R. S. (2018). *Modern Labor Economics: Theory And Public Policy* (13th Ed.). Routledge.

Ehrenberg, R. G., & Smith, R. S. (2018). Modern Labor Economics: Theory And Public Policy (13th Ed.). Routledge.

Gujarati, D. N., & Porter, D. C. (2017). *Basic Econometrics* (6th Ed.). Mcgraw-Hill Education. Basic Econometrics - Damodar N. Gujarati - Google Buku

Gujarati, D. N., & Porter, D. C. (2017). Basic Econometrics (6th Ed.). Mcgraw-Hill Education.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression (3rd Ed.). John Wiley & Sons.

Kharisma, B., & Wardhana, A. (2022). Determinants Of Labor Force Participation In Indonesia: A Spatial Panel Data Approach. International Journal Of Social Economics, 49(5), 721-738.

Kusuma, A. P., & Haryanto, T. (2024). Population Density And Economic Disparities: Impact On Regional Unemployment Rates In Emerging Economies. Journal Of Developing Areas, 58(1), 145-160.

Lestari, F., & Rahmawati, A. (2023). Binary Logistic Regression Modeling On Regional Poverty Classification In Central Java. Journal Of Applied Statistics And Data Science, 3(1), 22-34.

Mankiw, N. G. (2019). Macroeconomics (10th Ed.). Worth Publishers.

Nugroho, S., & Rahayu, T. (2023). Penerapan Generalized Linear Model (Glm) Pada Klasifikasi Kategori Pengangguran Daerah Di Indonesia. Jurnal Matematika Dan Aplikasi, 19(2), 78-89.

Pratama, A., & Wibowo, R. (2021). Analisis Pengaruh Kepadatan Penduduk Dan Tpak Terhadap Pengangguran Terbuka Di Jawa Barat. Jurnal Ilmu Ekonomi Terapan, 6(1), 55-67.

Pratama, Y., & Sukiyono, K. (2022). Influence Of Unemployment Rate And Labor Force Participation Rate On Regional Poverty Levels. Jurnal Ekonomi Dan Pembangunan Indonesia, 22(1), 45-58.

Putri, R. A., & Setiawan, I. (2023). Evaluasi Kebaikan Suai Dan Akurasi Klasifikasi Model Regresi Logistik Biner Pada Data Makroekonomi. Jurnal Aplikasi Statistik, 5(2), 112-124.

Rahayu, S., & Hartono, D. (2022). The Impact Of Human Development Index And Population Growth On Poverty Reduction: Evidence From Eastern Indonesia. Economics And Finance In Indonesia, 68(2), 89-101.

Rahman, M. F., & Suryadi, D. (2022). The Impact Of Human Development Index And Economic Growth On Labor Market Dynamics. Indonesian Journal Of Economics And Policy, 15(1), 40-53.

Sari, K. P., & Yulianto, E. (2021). Interpretasi Odds Ratio Pada Model Regresi Logistik Biner Untuk Pengelompokan Wilayah Rawan Pengangguran. Jurnal Statistika Dan Data Sains, 2(2), 88-97.

Siregar, H., & Amanda, D. (2022). Pengaruh Pdrb Per Kapita Dan Investasi Terhadap Penyerapan Tenaga Kerja Dan Tpt. Jurnal Ekonomi Pembangunan, 20(1), 15-28.

Suryanto, A., & Wibowo, R. (2023). Analisis Pengaruh Tpt Dan Tpak Terhadap Kemiskinan Regional Di Indonesia. Jurnal Ilmu Ekonomi Terapan, 8(2), 112-125.

Todaro, M. P., & Smith, S. C. (2020). Economic Development (13th Ed.). Pearson.

Utami, S., & Wahyudi, S. (2024). Uji Signifikansi Wald Dan Likelihood Ratio Test Pada Pemodelan Kategorik Biner Makroekonomi Regional. Jurnal Data Dan Statistik Terapan, 3(1), 12-25.

Wardhana, A., & Kharisma, B. (2021). Determinants Of Poverty In Indonesia: Empirical Evidence From Regional Panel Data Analysis. Journal Of Developing Areas, 55(3), 115-128.

Wati, L., & Kurniawan, H. (2023). Spatial Analysis Of Population Density And Labor Force Participation In Urban Areas. Asian Journal Of Regional Science, 7(2), 201-215.

Wijaya, C., & Santoso, B. (2022). Pemodelan Regresi Logistik Biner Pada Kategori Tingkat Pengangguran Terbuka Kabupaten/Kota Di Indonesia. Jurnal Matematika Dan Statistik, 8(3), 150-162.

Yulia, N., & Febriani, R. (2023). Analisis Pengaruh Ipm Dan Jumlah Penduduk Terhadap Tingkat Pengangguran Terbuka Menggunakan Approach Regresi Logistik. Jurnal Riset Ekonomi Dan Statistik, 3(4), 210-222.