

Klasifikasi Tingkat Kesejahteraan Kabupaten/Kota di Jawa Barat, Jawa Tengah, dan Jawa Timur Menggunakan Regresi Logistik Multinomial

Ahmad Firqi Nashrullah^{1,*}, I Nyoman Kresna Wira Yudha², Rivaldi Dwi Mahardhika³, Aviolla Terza Damaliana⁴, Shindi Shella May Wara⁵

^{1,2,3,4,5} Program Studi Sains Data, Universitas Pembangunan Nasional “Veteran” Jawa Timur, Jalan Raya Rungkut Madya, Kecamatan Gunung Anyar, Surabaya, 60294, Jawa Timur, Indonesia

*Corresponding author: 23083010009@student.upnjatim.ac.id



P-ISSN: 2986-4178
E-ISSN: 2988-4004

Riwayat Artikel

Dikirim: 13 Juni 2025
Direvisi: 01 Oktober 2025
Diterima: 01 Oktober 2025

ABSTRAK

Tingkat kesejahteraan daerah menjadi salah satu indikator utama dalam menilai kemajuan Pembangunan wilayah. Penelitian ini bertujuan untuk mengklasifikasikan kesejahteraan seluruh daerah setingkat kabupaten dan kota yang berada di wilayah Provinsi Jawa Barat, Jawa Tengah, serta Jawa Timur berdasarkan kelompok Indeks pembangunan manusia yang terdiri atas empat kategori, yaitu rendah, sedang, tinggi, dan sangat tinggi menggunakan regresi logistik multinomial. Analisis melibatkan persentase penduduk miskin, rasio ketimpangan, angka harapan hidup, pengeluaran per kapita, kepadatan penduduk, dan akses sanitasi layak. Data diperoleh dari Badan Pusat Statistik tahun 2023. Hasil deskriptif menunjukkan 69 wilayah termasuk kategori tinggi, 17 kategori sedang, dan 14 kategori sangat tinggi. Uji statistik mengonfirmasi hubungan signifikan semua variabel dan menunjukkan bahwa perbaikan akses sanitasi serta peningkatan harapan hidup meningkatkan indeks pembangunan manusia suatu wilayah. Model menghasilkan akurasi 86,67 persen. Hasil analisis ini dapat dimanfaatkan landasan objektif untuk menyusun strategi pembangunan wilayah yang efektif.

Kata Kunci: Indeks pembangunan manusia, Kesejahteraan wilayah, Sanitasi, Angka harapan hidup, Regresi logistik multinomial.

ABSTRACT

The level of regional welfare is a primary indicator of regional development progress. The objective of this study is to categorise the welfare of all district and city-level regions in the provinces of West Java, Central Java, and East Java based on the Human Development Index (HDI) group, which comprises four categories: low, medium, high, and very high. This categorisation will be achieved utilising multinomial logistic regression. The analysis encompassed a range of socio-economic indicators, including the percentage of the population living below the poverty line, the inequality ratio, life expectancy, per capita expenditure, population density, and access to proper sanitation. The data presented herein was obtained from the Central Bureau of Statistics in 2023. The descriptive results indicated that 69 regions were categorised as high, 17 as medium, and 14 as very high. The application of statistical tests confirmed the significant relationship between all variables, thereby demonstrating that enhanced access to sanitation and elevated life expectancy are positively correlated with an increase in the human development index of a region. The model demonstrated an accuracy of 86.67 per cent. The results of this analysis can be used as an objective basis for developing effective regional development strategies.

Keywords: Human development index, Regional welfare, Sanitation, Life expectancy, Multinomial logistic regression.

1. Pendahuluan

Kesejahteraan Masyarakat merupakan indikator utama dalam menilai keberhasilan pembangunan daerah [1]. Perbedaan tingkat kesejahteraan dan produktivitas antarwilayah di Pulau Jawa masih signifikan. Misalnya pada tahun 2023, Provinsi Jawa Barat (IPM 73,74), Provinsi Jawa Tengah (IPM 73,39), dan Jawa Timur (73,38) tercatat memiliki Indeks Pembangunan Manusia (IPM) terendah dibandingkan provinsi lain di Jawa. Artinya, kedua provinsi tersebut berada pada kategori tinggi berdasarkan kategori yang ditetapkan oleh Badan Pusat Statistik, yakni kurang dari 60 tergolong rendah, 60 hingga 69 termasuk sedang, 70 sampai 79 masuk kategori tinggi, dan 80 ke atas dianggap sangat tinggi [2], lebih rendah dari DKI Jakarta yang sudah mencapai kategori sangat tinggi (82,25 pada 2021) [3]. Di sisi lain, produktivitas tenaga kerja di Jateng dan Jatim relatif rendah jika dibandingkan dengan rata-rata nasional. Ketiga provinsi, yakni Jawa Barat, Jawa Tengah, dan Jawa Timur, yang merupakan wilayah dengan populasi terbesar di Indonesia, memiliki keragaman latar belakang dari segi sosial, ekonomi, maupun kondisi geografis. [4]. Ketiga provinsi ini memiliki jumlah kabupaten/kota terbanyak di Pulau Jawa serta tingkat keragaman penduduk yang tinggi. Data dari BPS menunjukkan adanya perbedaan mencolok dalam persentase penduduk miskin, ketimpangan pendapatan (rasio Gini), harapan hidup, pengeluaran per kapita, kepadatan penduduk, dan akses sanitasi layak diyakini memengaruhi kondisi kesejahteraan dan kapasitas produksi daerah [5], [6], [7]. Berbagai karakteristik daerah memiliki hubungan erat dengan capaian Indeks Pembangunan Manusia (IPM), yang menjadi tolak ukur utama kesejahteraan wilayah, sementara kemiskinan tinggi umumnya menurunkan IPM dan potensi ekonomi suatu daerah [8]. IPM dipengaruhi oleh beragam faktor seperti persentase penduduk miskin,

tingkat ketimpangan (rasio Gini), Angka Harapan Hidup (AHH), pengeluaran per kapita, kepadatan penduduk, dan akses rumah tangga terhadap sanitasi layak [9], [10]. Dalam penelitian ini, enam variabel digunakan sebagai variabel prediktor, yaitu persentase penduduk miskin, rasio gini, Angka Harapan Hidup (AHH), pengeluaran per kapita, kepadatan penduduk, dan akses sanitasi layak. Variabel-variabel tersebut dipandang relevan karena secara langsung atau tidak langsung memengaruhi tingkat capaian IPM yang menjadi variabel respon dalam model klasifikasi menggunakan regresi logistik multinomial.

Pemilihan regresi logistik multinomial berdasarkan pertimbangan yang didukung oleh sejumlah studi terdahulu karena telah terbukti efektif dalam mengkaji data pembangunan manusia di tingkat daerah. Analisis data IPM Provinsi Sumatera Barat tahun 2019 hingga 2021 menunjukkan bahwa angka harapan hidup menjadi variabel paling signifikan dalam model regresi logistik biner, dengan rasio odds sebesar 3,150 menandakan pengaruhnya terhadap klasifikasi IPM lebih besar dibandingkan tingkat pengangguran [11]. Penelitian terhadap kabupaten dan kota di Provinsi Jawa Barat dianalisis menerapkan regresi logistik biner dengan pendekatan seleksi variabel dan menghasilkan akurasi klasifikasi lebih dari 96,3 persen [12]. Sementara itu, analisis terhadap data tahun 2019 dari Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur di Pulau Jawa menunjukkan bahwa klasifikasi IPM berhasil dilakukan menggunakan regresi logistik ordinal dengan tingkat akurasi berkisar antara 73 hingga 83 persen, serta perbedaan faktor dominan antar ketiga provinsi tersebut [13].

Dalam penelitian kombinasi status IPKM dan IPM kabupaten/kota di Kalimantan dengan regresi logistik multinomial, parameter model diestimasi dengan metode *Maximum Likelihood* (MLE) dan pengujian hipotesis meliputi *Likelihood Ratio Test* (LRT) dan uji parsial (uji Wald) [14], sedangkan interpretasi dilakukan melalui rasio odds. Langkah-langkah serupa juga diuraikan dalam studi regresi logistik ordinal untuk data penyebaran COVID-19 di Indonesia tahun 2019, di mana *Likelihood Ratio Test* (LRT) digunakan untuk menguji signifikansi keseluruhan model, Wald untuk setiap koefisien, dan analisis odds ratio untuk interpretasi hasil [15]. Hasil-hasil tersebut mendukung penggunaan pendekatan regresi logistik multinomial dalam mengkaji pengaruh variabel-variabel tersebut terhadap peluang suatu daerah berada pada kategori IPM tertentu.

Tujuan dari penelitian ini adalah untuk mengklasifikasikan kesejahteraan wilayah kabupaten/kota di Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur berdasarkan karakteristik wilayah yang memengaruhi capaian Indeks Pembangunan Manusia (IPM). IPM dikategorikan ke dalam empat tingkatan menurut standar BPS, yaitu rendah, sedang, tinggi, dan sangat tinggi. Karakteristik wilayah yang dianalisis mencakup persentase penduduk miskin, rasio Gini, angka harapan hidup, pengeluaran per kapita, kepadatan penduduk, dan akses terhadap sanitasi layak. Seluruh variabel tersebut dipertimbangkan sebagai faktor utama yang membentuk profil kesejahteraan suatu daerah. Metode regresi logistik multinomial digunakan untuk mengetahui bagaimana kombinasi variabel-variabel tersebut memengaruhi kemungkinan suatu kabupaten/kota berada dalam kategori kesejahteraan tertentu. Hasil penelitian diharapkan dapat memberikan pemahaman mendalam mengenai pola kesejahteraan antardaerah serta mendukung perumusan kebijakan pembangunan yang berfokus pada peningkatan kualitas wilayah secara lebih tepat sasaran.

2. Metodologi Penelitian

2.1 Sumber Data

Data pada analisis diperoleh melalui studi dokumentasi yang dipublikasikan oleh

Badan Pusat Statistik (BPS) Jawa Barat, Jawa Tengah dan Jawa Timur 2023. Data dikumpulkan dari situs resmi BPS. Data memiliki 8 variabel, yaitu: Kabupaten/Kota, Persentase Penduduk Miskin (persen), Indeks Pembangunan Manusia (IPM), Rasio Gini, Angka Harapan Hidup (AHH), Pengeluaran per Kapita (Orang/Tahun), Kepadatan Penduduk per km persegi (Km2), dan Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak. Penelitian ini melibatkan 100 kabupaten/kota yang tersebar di provinsi Jawa Tengah, Jawa Barat, dan Jawa Timur.

2.2 Regresi Logistik Multinomial

Regresi logistik multinomial merupakan model pada machine learning yang digunakan untuk memprediksi variabel respon kategorikal nominal dengan lebih dari dua kelas yang tidak berurutan, model ini memproses hasil dengan tiga atau lebih kemungkinan kategori. Regresi logistik multinomial merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel respon (y) yang bersifat polikotomus atau multinomial [16] Berdasarkan [17] menjelaskan bahwa rumus dari regresi logistik multinomial:

$$P(Y = k) = \frac{\exp(\beta_{0,k} + \sum_{q=1}^Q \beta_{q,k} X_q)}{1 + \exp(\beta_{0,k} + \sum_{q=1}^Q \beta_{q,k} X_q)} \quad \text{Untuk } k = 2, \dots, k \quad (1)$$

$$P(Y = 1) = \frac{1}{1 + \exp(\beta_{0,k} + \sum_{q=1}^Q \beta_{q,k} X_q)} \quad \text{Untuk } k = 2, \dots, k \quad (2)$$

Keterangan:

- $P(Y = k)$: Probabilitas bahwa variabel respon Y berada pada kategori ke-k.
- $\beta_{0,k}$: Intersep (konstanta) untuk kategori ke-k.
- $\beta_{q,k}$: Koefisien regresi untuk prediktor ke-q pada kategori ke-k.
- X_q : Variabel prediktor ke-q.
- Q : Jumlah variabel prediktor.

Maximum Likelihood Estimation (MLE):

$$\ln = \left(\frac{P(Y = k)}{p(Y = 1)} \right) = \beta_{0,k} + \sum_{q=1}^Q \beta_{q,k} X_q \quad (3)$$

Keterangan:

- $\frac{p(Y=k)}{p(Y=1)}$: Odds atau perbandingan probabilitas suatu observasi berada pada kategori k dibanding kategori referensi 1.
- $\beta_{0,k} + \sum_{q=1}^Q \beta_{q,k} X_q$: Kombinasi linear dari prediktor.

Faktor Shrinkage

$$\ln = \left(\frac{P(Y = k)}{p(Y = 1)} \right) = \alpha_{0,k}^* + SMN_k \left(\sum_{q=1}^Q \beta_{q,k} X_q \right) \text{ untuk } k = 2, \dots, k \quad (4)$$

Keterangan:

- $\alpha_{0,k}^*$: Intersep yang telah disesuaikan.
- SMN_k : Faktor shrinkage (penyusutan) untuk mengurangi overfitting.

2.3 Uji Independensi

Uji independensi menggunakan chi-squared adalah uji statistik non-parametrik yang dipakai sebagai penentu apakah terjadi pengaruh yang relevan antara dua atau lebih variabel kategorikal yang ada pada penelitian. Uji Chi-Square dimanfaatkan untuk yang menguji independensi pada dua variabel kategori [18]. Tujuan dari uji Chi Square adalah digunakan sebagai menguji hipotesis bahwasannya tidak terdapat pengaruh antara populasi atau kriteria, dua atau lebih kelompok, dan untuk menguji sejauh mana distribusi data yang diamati sesuai dalam distribusi yang diharapkan [19].

$$X^2 = \sum_{i=0}^j \frac{(O_i - E_i)^2}{E_i} \quad (5)$$

Keterangan:

- E_i : Jumlah frekuensi yang diharapkan untuk tingkat ke-I dari variabel kategori.
- O_i : Jumlah frekuensi yang diamati untuk tingkat ke-i dari variabel kategori.

Derajat bebas (degree of freedom) pada uji independensi dihitung dengan rumus $(r - 1)(c - 1)$, dimana r adalah jumlah kolom dalam tabel dan c adalah jumlah baris pada tabel. Jika hasil pengujian menunjukkan nilai chi-squared yang signifikan, maka dapat disimpulkan bahwa terdapat hubungan antara kedua variabel yang diuji [19].

2.4 Uji Parameter Secara Serentak (Uji G)

Uji parameter secara serentak merupakan uji signifikansi sekelompok parameter secara bersama-sama dalam sebuah model statistik berbasis likelihood. Likelihood Ratio Test (LRT) digunakan untuk mengevaluasi kesesuaian model dengan variabel prediktor secara keseluruhan [15]. Uji serentak digunakan sebagai uji pengaruh variabel bebas pada model regresi secara keseluruhan [20]. Uji signifikan parameter secara serentak menggunakan hipotesis:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{Minimal terdapat satu } \beta_j \neq 0, \text{ dengan } j = 1, 2, \dots, p$$

Menggunakan rumus uji statistik menurut [21].

$$G = -2 \ln \left[\frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_2}{n}\right)^{n_2} \left(\frac{n_3}{n}\right)^{n_3}}{\prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}}]} \right] \quad (6)$$

Keterangan:

- $\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_2}{n}\right)^{n_2} \left(\frac{n_3}{n}\right)^{n_3}$: likelihood dari model independen (model tanpa prediktor).
- n : Total jumlah observasi.
- $\pi_j(x_i)$: probabilitas bahwa observasi ke-iii masuk kategori jjj yang diprediksi oleh model.
- y_{ji} : indikator (0 atau 1), bernilai 1 jika observasi ke-iii termasuk kategori jjj, dan 0 jika tidak.

Kriteria uji:

Tolak H_0 jika $G > X_{(d,b,a)}^2$ atau P-value $< \alpha$ (0,05)

2.5 Uji Parsial (Uji Wald)

Sebuah Metode statistik yang dimanfaatkan sebagai yang menguji seberapa signifikansi parameter tunggal (koefisien) dalam model regresi atau model statistik multivariat, sementara variabel lain dianggap konstan. Uji parsial dilakukan sebagai yang menilai partisipasi atau tingkat signifikansi setiap variabel independen terhadap model yang digunakan [22]. Uji signifikan wald menggunakan hipotesis:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 ; j = 1, 2, \dots, p$$

Statistik uji yang dipakai pada uji parsial wald menurut [21].

$$W = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (7)$$

Keterangan:

- $\hat{\beta}_j$: Penaksir dari β_j .
- $SE(\hat{\beta}_j)$: Penaksir galat baku β_j .

Kriteria uji:

Tolak H_0 jika $W > X_{(d,b,a)}^2$ atau P-value $< \alpha$ (0,05)

2.6 Odds Ratio

Odds Ratio (OR) merupakan ukuran statistik yang digunakan untuk mengevaluasi kekuatan hubungan antara variabel bebas terhadap faktor risiko dengan kejadian tertentu, seperti penyakit, kondisi sosial, atau status tertentu. OR menunjukkan rasio probabilitas terjadinya suatu hasil tertentu pada kelompok yang terpapar dibandingkan dengan kelompok yang tidak terpapar [23]. Rasio peluang (odds ratio) dilambangkan dengan OR, yang didefinisikan sebagai perbandingan antara nilai peluang untuk $x = 1$ dengan nilai peluang untuk $x = 0$, dan persamaan berikut dapat dibentuk [15]. Rumus yang digunakan untuk menentukan odds ratio ketika menggunakan model regresi logistik, yaitu [15]:

$$OR = \frac{\frac{\pi(1)}{[1 - \pi(1)]}}{\frac{\pi(0)}{[1 - \pi(0)]}} = \frac{\frac{e^{\beta_0 + \beta_1}}{1 + e^{\beta_0 + \beta_1}}}{\frac{1}{1 + e^{\beta_0}}} = e^{(\beta_0 + \beta_1) - \beta_0} \quad (8)$$

Keterangan:

- $\pi(1)$: Probabilitas Y = 1 ketika X = 1.
- $\pi(0)$: Probabilitas Y = 1 ketika X = 0.
- $\frac{e^{\beta_0 + \beta_1}}{e^{\beta_0}} = e^{\beta_1}$: Odds Ratio.

Dengan interpretasi nilai OR:

1. $OR > 1$ menandakan adanya efek karena nilai probabilitas keberhasilan lebih besar daripada nilai probabilitas kegagalan.
2. $OR = 1$ menunjukkan tidak adanya efek karena nilai probabilitas keberhasilan sama dengan probabilitas kegagalan.
3. $OR < 1$ menandakan adanya probabilitas keberhasilan lebih kecil daripada nilai probabilitas kegagalan.

2.7 Uji Kesesuaian Model (Goodness of Fit)

Pengujian kesesuaian model adalah suatu prosedur statistik yang digunakan untuk mengevaluasi seberapa baik suatu model statistika menjelaskan data observasi. Model terbaik didapatkan dengan mempertimbangkan nilai kriteria Informasi Akaike (AIC) dengan menghitung jumlah parameter, semakin kecil nilai Kriteria Informasi Akaike (AIC), semakin baik kualitas modelnya [22]. Tujuan utamanya untuk memastikan bahwa prediksi model cocok dengan data sebenarnya atau terdapat ketidaksesuaian sistematis antara model dan kenyataan. Dalam analisis data, diperlukan untuk menentukan derajat bebas (df), jumlah frekuensi yang diharapkan, statistik uji dan p-value yang terkait dengan statistik uji [19]. $df = k - 1$, k = jumlah kategori data sampel. Dengan menggunakan rumus menurut [24].

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (9)$$

Dengan $\bar{\pi}_k = \left(\sum_{j=1}^{c_k} \frac{m_j \hat{\pi}_j}{n'_k} \right)$

Keterangan :

- O_k : Observasi grup ke-k.
- $\bar{\pi}_k$: Estimasi Probabilitas rata-rata.
- g : Jumlah grup (kombinasi kategori dalam model serentak).
- n'_k : Banyak observasi pada grup ke-k.

3. Hasil dan Pembahasan

3.1. Analisis Deskriptif

Penelitian ini menggunakan 100 observasi dari kabupaten/kota di Provinsi Jawa Barat, Jawa Tengah, Jawa Timur yang bertujuan untuk menganalisis faktor-faktor penentu terhadap Indeks Pembangunan Manusia (IPM). IPM merupakan indikator komposit yang mencerminkan kualitas hidup masyarakat dari aspek pendidikan, kesehatan, dan ekonomi. Berdasarkan hasil klasifikasi, data menunjukkan bahwa 69% daerah tergolong dalam kategori Tinggi, 17% dalam kategori Sedang, dan hanya 14% dalam kategori Sangat Tinggi. Hal ini mencerminkan ketimpangan pembangunan antarwilayah masih terjadi, meskipun secara umum, sebagian besar wilayah berada dalam status pembangunan manusia yang cukup baik.

Tabel 1 Hasil Statistika Deskriptif

Variabel	Mean	Median	Std Dev	Min	Max
Persentase Penduduk Miskin (persen)	0,06	-0,04	0,99	-1,76	3,34
Rasio Gini	-0,03	-0,21	0,99	-2,68	2,51
Angka Harapan Hidup (AHH)	-0,04	-0,10	1,06	-3,85	2,68
Pengeluaran per Kapita (Orang/Tahun)	-0,02	-0,33	1,01	-1,55	3,28
Kepadatan Penduduk per km persegi (Km ²)	0,03	-0,45	1,04	-0,67	3,95
Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	-0,02	0,22	1,06	-2,72	1,18

Berdasarkan 100 data observasi dari kabupaten/kota di Jawa Barat, Jawa Tengah, dan Jawa Timur, hasil statistik deskriptif terhadap enam variabel predictor yang telah terstandarisasi menunjukkan bahwa seluruh variabel memiliki distribusi nilai yang wajar. Nilai rata-rata yang mendekati nol serta standar deviasi yang berada di kisaran satu mengindikasikan bahwa proses standarisasi data telah berhasil dilakukan dengan baik.

3.2. Analisis Regresi Logistik Multinomial

3.2.1. Uji Independensi

Sebelum membangun model regresi logistik multinomial, dilakukan Uji Chi-Squared untuk Independensi guna menilai apakah terdapat hubungan yang signifikan antara masing-masing variabel prediktor dengan kategori Indeks Pembangunan Manusia (IPM). Karena variabel prediktor bersifat numerik, maka terlebih dahulu dilakukan proses pengelompokan (*binning*) menjadi tiga kategori yaitu rendah, sedang, tinggi, dan rendah (tidak muncul karena dalam dataset tidak ada daerah yang masuk dalam kategori rendah) dengan menggunakan teknik *quantile discretization* atau *equal-frequency discretization* merupakan metode pengelompokan data numerik dengan membaginya ke dalam sejumlah interval (*bin*) berdasarkan kuantil distribusi data. Dengan cara ini, setiap interval akan memiliki jumlah data yang relatif sama meskipun rentang nilainya bisa berbeda. Metode ini efektif digunakan untuk menjaga keseimbangan jumlah observasi dalam setiap kategori, terutama ketika distribusi data tidak merata, dan dinilai lebih baik dibandingkan metode

equal-width discretization yang hanya membagi data berdasarkan lebar interval yang sama [24]. Uji dilakukan pada tingkat signifikansi $\alpha = 0,05$, yang berarti hipotesis nol (H_0 : tidak ada hubungan antara variabel) akan ditolak jika nilai p-value $< 0,05$.

Tabel 2 Hasil Uji Independen Chi-Squared

Variabel	Chi-Square	P-Value	Keterangan
Persentase Penduduk Miskin (persen)	30,7521	0,0000	Terdapat hubungan signifikan dengan kategori IPM
Rasio Gini	17,7785	0.0014	Terdapat hubungan signifikan dengan kategori IPM
Angka Harapan Hidup (AHH)	31,2184	0,0000	Terdapat hubungan signifikan dengan kategori IPM
Pengeluaran per Kapita (Orang/Tahun)	29,2089	0,0000	Terdapat hubungan signifikan dengan kategori IPM
Kepadatan Penduduk per km persegi (Km ²)	29,8430	0,0000	Terdapat hubungan signifikan dengan kategori IPM
Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	21,1326	0,0003	Terdapat hubungan signifikan dengan kategori IPM

Hasil pengujian menunjukkan bahwa seluruh variabel prediktor (independen) memiliki p-value $< 0,05$, yang menandakan adanya hubungan signifikan secara statistik dengan klasifikasi IPM; oleh karena itu, semua variabel tersebut layak digunakan dalam pemodelan regresi logistik multinomial guna memastikan tidak adanya variabel bebas yang tidak relevan serta mendukung keakuratan model yang akan dibentuk.

3.2.2. Uji Parameter secara Serentak (Uji G)

Uji ini dilakukan untuk menguji hipotesis nol bahwa semua koefisien regresi sama dengan nol. Hasil menunjukkan nilai G-statistic sebesar 34.000 dengan derajat bebas 6 dan p-value sebesar 0.0000. Karena p-value < 0.05 , maka H_0 ditolak, yang berarti model secara simultan signifikan. Ini menunjukkan bahwa paling tidak ada satu variabel prediktor yang signifikan dalam memengaruhi klasifikasi kategori IPM.

3.2.3. Uji Parsial (Uji Wald)

Uji parsial (uji Wald) dilakukan untuk mengevaluasi pengaruh masing-masing variabel independen terhadap peluang suatu wilayah berada dalam kategori tertentu dari IPM. Dengan kata lain, uji ini menguji hipotesis nol bahwa setiap koefisien parameter regresi sama dengan nol ($H_0: \beta = 0$) terhadap alternatif ($H_1: \beta \neq 0$), pada tingkat signifikansi $\alpha = 0,05$. Jika nilai p-value dari suatu parameter $< 0,05$, maka dapat disimpulkan bahwa variabel tersebut secara signifikan memengaruhi klasifikasi IPM. Hasil koefisien dari model dikonversi ke dalam bentuk Odds Ratio (OR) untuk memudahkan interpretasi.

Tabel 3 Asumsi Nilai Variabel

Variabel	Notasi	Asumsi Nilai	Standar Deviasi	Keterangan
Persentase Penduduk Miskin (persen)	X1	1	0,99	Peningkatan 1 standar deviasi kemiskinan
Rasio Gini	X2	1	0,99	Peningkatan 1 standar deviasi ketimpangan
Angka Harapan Hidup (AHH)	X3	1	1,06	Peningkatan harapan hidup
Pengeluaran per Kapita (Orang/Tahun)	X4	1	1,01	Peningkatan daya beli masyarakat
Kepadatan Penduduk per km persegi (Km2)	X5	1	1,04	Wilayah lebih padat
Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	X6	1	1,06	Peningkatan akses sanitasi

Seluruh variabel prediktor dalam penelitian ini telah distandarisasi menggunakan metode z-score, yaitu dengan mengurangi nilai rata-rata dan membaginya dengan standar deviasi dari masing-masing variabel. Hasil standarisasi menghasilkan nilai standar deviasi yang mendekati 1, yakni berkisar antara 0,99 hingga 1,06, yang menandakan bahwa proses normalisasi telah dilakukan dengan baik. Oleh karena itu, dalam interpretasi hasil uji Wald dan perhitungan Odds Ratio (OR), digunakan asumsi bahwa setiap variabel meningkat sebesar satu satuan skor-z ($X = 1$), atau setara dengan peningkatan ± 1 standar deviasi dari rata-rata variabel aslinya. Pendekatan ini bertujuan untuk menjaga konsistensi antarvariabel dan memungkinkan pembacaan langsung pengaruh β terhadap log-odds tanpa perlu transformasi tambahan. Dengan menggunakan asumsi $X = 1$, nilai koefisien β dapat dikonversi langsung ke OR melalui persamaan $OR = \exp(\beta)$, yang menggambarkan seberapa besar peluang perubahan kategori IPM akibat peningkatan satu unit standar deviasi dari suatu variabel, dengan variabel lainnya diasumsikan konstan. Pendekatan ini sah dan umum digunakan dalam model klasifikasi berbasis regresi logistik yang melibatkan data terstandarisasi, serta memperkuat validitas komparatif antarvariabel dalam model multinomial.

Tabel 4 Hasil Uji Parsial (Uji Wald) Kategori IPM Sedang

Variabel (IPM Sedang)	Koefisien (β)	Odds Ratio (OR)	P-Value	Keterangan
Persentase Penduduk Miskin (persen)	-1,1447	0,318	0,0000	Meningkatnya kemiskinan menurunkan peluang masuk ke kategori IPM Sedang.
Rasio Gini	0,8729	2,393	0,0000	Kenaikan Rasio Gini meningkatkan peluang masuk ke kategori IPM Sedang.
Angka Harapan Hidup (AHH)	1,1594	3,189	0,0000	Wilayah dengan AHH lebih tinggi cenderung masuk kategori IPM Sedang.
Pengeluaran per Kapita (Orang/Tahun)	1,4881	4,428	0,0000	Peningkatan daya beli meningkatkan peluang IPM Sedang secara signifikan.
Kepadatan Penduduk per km persegi (Km2)	0,7572	2,132	0,0016	Wilayah padat penduduk cenderung masuk ke kategori IPM Sedang.
Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	1,2799	3,596	0,0000	Akses sanitasi yang baik meningkatkan peluang masuk kategori IPM Sedang.

Pada kategori IPM Sedang, mayoritas variabel independen memberikan kontribusi positif terhadap peluang masuk ke kelas ini. Misalnya, peningkatan akses terhadap sanitasi layak meningkatkan kemungkinan suatu wilayah diklasifikasikan dalam kategori ini lebih dari tiga kali lipat (OR = 3,596). Demikian pula, pengeluaran per kapita (OR = 4.428), angka harapan hidup (OR = 3.189), dan kepadatan penduduk (OR = 2.132) juga memperbesar peluang secara signifikan. Artinya, wilayah dengan kondisi sosial-ekonomi dan infrastruktur yang lebih baik cenderung memiliki IPM yang berada di level sedang. Satu-satunya variabel dengan pengaruh negatif adalah persentase penduduk miskin, yang menunjukkan bahwa peningkatan kemiskinan secara nyata menurunkan kemungkinan suatu daerah tergolong dalam kelompok IPM sedang. Sebagian besar variabel independen menunjukkan pengaruh positif dan signifikan terhadap peluang suatu wilayah berada dalam kategori IPM Sedang yang dapat dilihat dari fungsi logit berikut.

$$g(\text{Sedang}) = -1,1447X_1 + 0,8729X_2 + 1,1594X_3 + 1,4881X_4 + 0,7572X_5 + 1,2799X_6$$

$$g(\text{Sedang}) = -1,1447 + 0,8729 + 1,1594 + 1,4881 + 0,7572 + 1,2799 = 4,4128$$

$$P(\text{Sedang}) = \frac{e^{4,4128}}{1 + e^{4,4128} + e^{-4,5535} + e^{0,1407}} \approx 0,986$$

Dengan asumsi bahwa semua variabel meningkat sebesar 1 standar deviasi, maka log-odds suatu wilayah untuk diklasifikasikan ke dalam kategori IPM Sedang adalah 4,41, yang berarti peluangnya cukup besar. Hal ini menunjukkan bahwa wilayah dengan nilai sosial ekonomi dan infrastruktur dasar yang relatif meningkat (misalnya AHH, pengeluaran, sanitasi) cenderung diklasifikasikan dalam IPM Sedang secara signifikan, selama belum melampaui ambang batas menuju kategori yang lebih tinggi.

Tabel 5 Hasil Uji Parsial (Uji Wald) Kategori IPM Tinggi

Variabel (IPM Tinggi)	Koefisien (β)	Odds Ratio (OR)	P-Value	Keterangan
Persentase Penduduk Miskin (persen)	1,0845	2,958	0,0000	Meningkatnya kemiskinan menaikkan peluang masuk ke kategori IPM Tinggi.
Rasio Gini	-0,4232	0,655	0,1638	Kenaikan Rasio Gini menurunkan peluang masuk ke kategori IPM Tinggi.
Angka Harapan Hidup (AHH)	-1,4982	0,224	0,0000	Harapan hidup tinggi menurunkan peluang masuk kategori IPM Tinggi.
Pengeluaran per Kapita (Orang/Tahun)	-1,1321	0,322	0,0001	Semakin tinggi pengeluaran, semakin kecil peluang masuk kategori IPM Tinggi.
Kepadatan Penduduk per persegi (Km2)	-1,0421	0,353	0,0000	Peningkatan kepadatan penduduk menurunkan peluang wilayah masuk ke kategori IPM Tinggi.
Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	-1,5424	0,213	0,0000	Akses sanitasi yang lebih baik mengurangi kemungkinan wilayah berada di kategori IPM Tinggi.

Hasil pada kategori IPM Tinggi memperlihatkan bahwa seluruh variabel justru memiliki pengaruh negatif terhadap peluang masuk ke dalam kelas ini. Hal ini mengindikasikan bahwa ketika suatu wilayah menunjukkan peningkatan dalam indikator-indikator seperti harapan hidup, pengeluaran per kapita, dan akses sanitasi, maka kemungkinan besar wilayah tersebut melewati ambang IPM Tinggi dan masuk ke kategori yang lebih tinggi, yaitu IPM Sangat Tinggi. Sebaliknya, wilayah dengan indikator pembangunan yang lebih rendah masih memiliki kemungkinan berada di kategori ini. Fakta bahwa kemiskinan justru berkontribusi positif di sini (OR = 2,958) menandakan bahwa kategori ini bisa jadi berperan sebagai zona transisi, di mana wilayah yang sedang berkembang menunjukkan pergeseran dari status sedang menuju sangat tinggi. Hal ini tercermin dari nilai odds ratio yang rendah (OR < 1) pada sebagian besar indikator pembangunan, menunjukkan bahwa peningkatan nilai indikator justru mengurangi kemungkinan wilayah tertahan dalam kategori ini. Pada kategori ini, sebagian besar variabel justru memiliki pengaruh negatif dan signifikan, kecuali persentase penduduk miskin, fungsi yang terbentuk sebagai berikut.

$$g(Tinggi) = 1,0845X_1 - 0,4232X_2 - 1,4982X_3 - 1,1321X_4 - 1,0421X_5 - 1,5424X_6$$

$$g(Tinggi) = -1,0845 - 0,4232 - 1,4982 - 1,1321 - 1,0421 - 1,5424 = -4,5535$$

$$P(Tinggi) = \frac{e^{-4,5535}}{1 + e^{4,4128} + e^{-4,5535} + e^{0,1407}} \approx 0,0001$$

Log-odds sebesar -4,55 menunjukkan bahwa dengan peningkatan serempak semua indikator, peluang suatu wilayah berada pada kategori IPM Tinggi justru menurun drastis.

Ini mendukung argumen bahwa wilayah dengan performa pembangunan tinggi (misalnya AHH, pengeluaran, dan sanitasi) akan melampaui kategori IPM Tinggi dan masuk ke level lebih tinggi, yaitu IPM Sangat Tinggi. Sebaliknya, jika nilai-nilai variabel tersebut masih rendah, peluang untuk berada di kategori IPM Tinggi lebih besar.

Tabel 6 Hasil Uji Parsial (Uji Wald) Kategori IPM Sangat Tinggi

Variabel (IPM Sangat Tinggi)	Koefisien (β)	Odds Ratio (OR)	P-Value	Keterangan
Persentase Penduduk Miskin (persen)	0,0602	1,062	0,7651	Hampir tidak berpengaruh; sedikit meningkatkan peluang kategori IPM Sangat Tinggi.
Rasio Gini	-0,4497	0,638	0,1170	Kenaikan Rasio Gini menurunkan peluang masuk ke kategori IPM Sangat Tinggi.
Angka Harapan Hidup (AHH)	0,3388	1,403	0,0999	Peningkatan AHH sedikit meningkatkan peluang masuk kategori IPM Sangat Tinggi.
Pengeluaran per Kapita (Orang/Tahun)	-0,3560	0,700	0,1666	Pengaruh negatif ringan, tidak dominan terhadap peluang kategori IPM Sangat Tinggi.
Kepadatan Penduduk per km persegi (Km ²)	0,2848	1,329	0,2679	Wilayah padat cenderung memiliki peluang yang sedikit lebih tinggi untuk masuk ke kategori IPM Sangat Tinggi.
Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak	0,2626	1,300	0,2074	Akses sanitasi yang baik sedikit meningkatkan peluang masuk kategori IPM Sangat Tinggi.

Untuk kategori IPM Sangat Tinggi, sebagian besar variabel memperlihatkan arah pengaruh positif, namun tidak terlalu kuat. Angka harapan hidup, kepadatan penduduk, dan akses terhadap sanitasi layak menunjukkan kontribusi positif terhadap peluang suatu daerah tergolong dalam kategori ini, dengan nilai Odds Ratio berkisar antara 1,3 hingga 1,4. Hal ini menunjukkan bahwa wilayah yang masuk kategori ini cenderung telah mencapai standar infrastruktur dan layanan dasar yang cukup baik. Di sisi lain, variabel seperti rasio Gini dan pengeluaran per kapita justru memberikan efek negatif ringan. Ini dapat diartikan bahwa ketimpangan dan faktor ekonomi mungkin sudah tidak lagi menjadi pembeda utama pada level kesejahteraan tertinggi. Persentase penduduk miskin memiliki pengaruh sangat lemah (OR = 1), menandakan dampaknya hampir netral dalam membedakan kategori ini. Untuk fungsi dari kategori ini sebagai berikut.

$$g(\text{Sangat Tinggi}) = 0,0602X_1 - 0,4497X_2 + 0,3388X_3 - 0,3560X_4 + 0,2848X_5 + 0,2626X_6$$

$$g(\text{Sangat Tinggi}) = 0,0602 - 0,4497 + 0,3388 - 0,3560 + 0,2848 + 0,2626 = 0,1407$$

$$P(\text{Sangat Tinggi}) = \frac{e^{0,1407}}{1 + e^{4,4128} + e^{-4,5535} + e^{0,1407}} \approx 0,0137$$

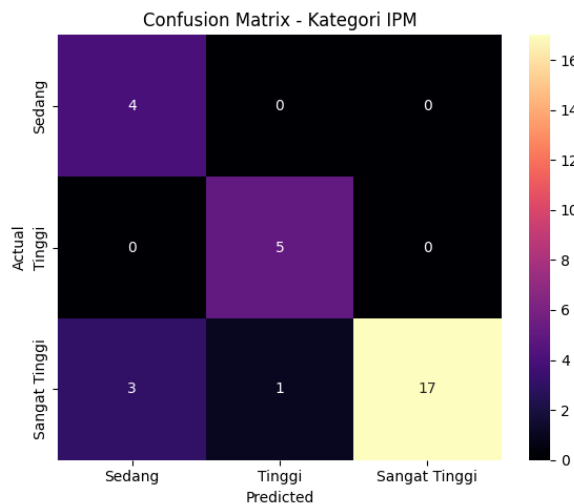
Log-odds sebesar 0,14 mengindikasikan peluang yang relatif seimbang bagi suatu wilayah untuk diklasifikasikan sebagai IPM Sangat Tinggi. Artinya, meskipun semua indikator meningkat satu standar deviasi, pengaruhnya tidak sepenuhnya kuat untuk mendorong wilayah ke kategori tertinggi, namun sudah menunjukkan tanda-tanda transisi menuju pembangunan manusia yang sangat baik. Berdasarkan hasil perhitungan fungsi logit, diketahui bahwa wilayah yang mengalami peningkatan pada seluruh variabel prediktor sebesar satu standar deviasi memiliki peluang sebesar 98,6% untuk diklasifikasikan ke dalam kategori IPM Sedang, sementara hanya 0,01% kemungkinan masuk ke kategori IPM Tinggi, dan sekitar 1,37% berada dalam kategori IPM Sangat Tinggi. Temuan ini menunjukkan bahwa meskipun terdapat peningkatan serempak pada faktor-faktor sosial ekonomi dan infrastruktur dasar, sebagian besar wilayah masih tertahan dalam kategori IPM Sedang. Hal ini mengindikasikan bahwa diperlukan lompatan yang lebih besar dan terintegrasi pada indikator-indikator kunci agar suatu daerah dapat bergerak menuju kategori pembangunan manusia yang lebih tinggi.

3.2.4. Uji Kesusaian Model (Goodness of Fit)

Tujuan dari pemodelan regresi logistik multinomial dalam penelitian ini adalah untuk membangun model yang mampu mengklasifikasikan kategori variabel dependen secara akurat dan representatif terhadap data yang diamati. Melalui uji Goodness of Fit, dievaluasi sejauh mana model dapat menyesuaikan diri dengan data aktual, di mana semakin tinggi tingkat akurasi dan konsistensi antara nilai prediksi dengan nilai observasi, maka semakin baik pula kemampuan model dalam merepresentasikan struktur data yang sesungguhnya.

Tabel 7 Hasil Uji Goodness of Fit

Uji Goodness of Fit	Score
Akurasi Model	0,8667
Score Model	0,8667



Gambar 1 Confusion Matrix

Hasil pengujian menunjukkan bahwa model memiliki akurasi sebesar 86.67%, yang artinya lebih dari 86% data uji dapat diprediksi dengan benar oleh model. Selain itu, nilai score model yang juga sebesar 0.8667 menegaskan bahwa performa prediktif model sangat konsisten. Berdasarkan confusion matrix, klasifikasi paling akurat terjadi pada kategori "Tinggi", disusul oleh "Sedang" dan "Sangat Tinggi". Temuan ini memperkuat

bahwa model regresi logistik multinomial yang dikembangkan memiliki validitas statistik yang baik untuk digunakan dalam mengklasifikasikan status pembangunan manusia suatu wilayah berdasarkan variabel-variabel yang dianalisis.

4. Kesimpulan

Penelitian ini menyimpulkan bahwa model regresi logistik multinomial secara signifikan mampu mengklasifikasikan tingkat kesejahteraan kabupaten/kota di Jawa Barat, Jawa Tengah, dan Jawa Timur berdasarkan enam indikator sosial-ekonomi. Mayoritas wilayah tergolong dalam kategori kesejahteraan tinggi, meskipun masih terdapat ketimpangan dengan adanya daerah yang masuk dalam kategori kesejahteraan sedang dan sangat tinggi. Semua variabel prediktor menunjukkan hubungan bermakna dengan klasifikasi kesejahteraan menurut uji Chi-Square, dan kelayakan model terkonfirmasi melalui uji G dengan tingkat akurasi mencapai 86,67%. Analisis parsial (uji Wald) memperlihatkan bahwa setiap kategori tingkat kesejahteraan dipengaruhi oleh variabel yang berbeda: daerah berkesejahteraan sedang dipengaruhi oleh faktor-faktor positif seperti akses air bersih dan sanitasi, sedangkan daerah berkesejahteraan tinggi menunjukkan gejala transisi dengan efek negatif dari sebagian besar variabel, dan kategori sangat tinggi ditentukan oleh keseimbangan kontribusi variabel-variabel tersebut. Secara keseluruhan, model ini memberikan landasan yang kokoh bagi perumusan kebijakan pembangunan yang berbasis data untuk meningkatkan kesejahteraan masyarakat.

5. Daftar Pustaka

- [1] R. Sihite, "ANALISIS PENGARUH PENDAPATAN PERKAPITA, JUMLAH KONSUMSI DAN PERTUMBUHAN EKONOMI DI KABUPATEN/ KOTA PROVINSI KALIMANTAN TENGAH," *J. Ekon. Pembangunan dan Bisnis*, vol. 2, no. April, pp. 46–57, 2022.
- [2] BPS Indonesia, "Indeks Pembangunan manusia (Metode Baru)," *Badan Pus. Stat. Indones.*, vol. 2015, no. September, p. 90, 2015, [Online]. Available: <https://www.bps.go.id/linkTableDinamis/view/id/1211>
- [3] Badan Pusat Statistik, "Indeks Pembangunan Manusia (IPM) 2023," *Badan Pus. Stat.*, no. 80, pp. 1–12, 2023, [Online]. Available: <https://www.bps.go.id/id/pressrelease/2023/11/15/2023/indeks-pembangunan-manusia--ipm--indonesia-tahun-2023-mencapai-74-39--meningkat-0-62-poin--0-84-persen--dibandingkan-tahun-sebelumnya--73-77--.html>
- [4] "Penduduk, Laju Pertumbuhan Penduduk, Distribusi Persentase Penduduk, Kepadatan Penduduk, Rasio Jenis Kelamin Penduduk Menurut Provinsi, 2023," 2024. [Online]. Available: <https://www.bps.go.id/id/statistics-table/3/V1ZSbFRUY3ITbFpEYTNsVWNGcDZjek53YkhsNFFUMDkjMyMwMDAw/jumlah-penduduk--laju-pertumbuhan-penduduk--distribusi-persentase-penduduk--kepadatan-penduduk--rasio-jenis-kelamin-penduduk-menurut-provinsi.html?year=2023>
- [5] S. Fauziyah, "Statistik Kesejahteraan Rakyat Provinsi Jawa Tengah 2023," *BPS Jawa Teng.*, vol. 9, no. 1, pp. 1–374, 2023, [Online]. Available: http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PEMBETUNGAN_TERPUSAT_STRATEGI_MELESTARI
- [6] R. Wulandari, "Statistik Kesejahteraan Rakyat Provinsi Jawa Barat 2023," *BPS Jawa Barat*, vol. 9, no. 1, pp. 1–252, 2023, [Online]. Available: http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PEMBETUNGAN_TERPUSAT_STRATEGI_MELESTARI

AN_TERPUSAT_STRATEGI_MELESTARI

- [7] E. Cahyono, N. Wulan, V. Paramita, and A. Yufinanda, "Statistik Kesejahteraan Rakyat Provinsi Jawa Timur 2023," *Badan Pus. Stat. Provinsi Jawa Timur*, vol. 6, pp. 1–453, 2023.
- [8] P. Ekonomi and U. N. Surabaya, "Determinan Indeks Pembangunan Manusia di Jawa Tengah Periode 2015-2023," vol. 4, no. 3, pp. 5845–5855, 2025.
- [9] J. V. Simanjuntak, "Pengaruh Sanitasi, Air Minum, Dan Dana Alokasi Khusus (DAK) Fisik Reguler Terhadap Indeks Pembangunan Manusia," *J. Law, Adm. Soc. Sci.*, vol. 4, no. 2, pp. 317–327, 2024, doi: 10.54957/jolas.v4i2.690.
- [10] E. D. A. Chaniago and M. Hanri, "Efek Aglomerasi di Indonesia : Komparasi Jawa dan Luar Jawa," *J. Kebijak. Ekon.*, vol. 16, no. 2, p. 9, 2021.
- [11] F. R. Ul Hasanah and M. Yollanda, "Penerapan Model Regresi Logistik Terhadap Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Barat Tahun 2019 – 2021," *JOSTECH J. Sci. Technol.*, vol. 2, no. 2, pp. 199–208, 2022, doi: 10.15548/jostech.v2i2.4383.
- [12] H. Setiawan, "ANALISIS REGRESI LOGISTIK BINER MENGGUNAKAN METODE SPARSE-GROUP LASSO PADA DATA BERDIMENSI TINGGI," vol. April, pp. 1–23, 2025.
- [13] M. Sari and P. Purhadi, "Pemodelan Indeks Pembangunan Manusia Provinsi Jawa Barat, Jawa Timur Dan Jawa Tengah Tahun 2019 Dengan Menggunakan Metode Regresi Logistik Ordinal," *J. Gaussian*, vol. 10, no. 1, pp. 149–158, 2021, doi: 10.14710/j.gauss.v10i1.30022.
- [14] Y. Paliling, M. Fathurahman, and S. Wahyuningsih, "Multinomial Logistic Regression To Model The Combination Of Phdi Status And Hdi Status Of Districts/Cities In Kalimantan Island," *J. Mat. Stat. dan Komputasi*, vol. 19, no. 3, pp. 460–472, 2023, doi: 10.20956/j.v19i3.22299.
- [15] T. M. Dewi and R. Kusumawati, "an Ordinal Logistic Regression Model for Analyzing Risk Zone Status of Covid-19 Spread," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 16, no. 3, pp. 853–860, 2022, doi: 10.30598/barekengvol16iss3pp853-860.
- [16] Z. Z. Zahroh and I. Zain, "Analisis Regresi Logistik Multinomial Pada Faktor-Faktor Yang Mempengaruhi Sumber Air Bersih Rumah Tangga Di Jawa Timur," *J. Sains dan Seni ITS*, vol. 7, no. 2, 2019, doi: 10.12962/j23373520.v7i2.34701.
- [17] A. Pate *et al.*, "Minimum sample size for developing a multivariable prediction model using multinomial logistic regression," *Stat. Methods Med. Res.*, vol. 32, no. 3, pp. 555–571, 2023, doi: 10.1177/09622802231151220.
- [18] U. Maryam, W. Somayasa, R. Ruslan, L. G. La Gubu, and J. Jufra, "Estimasi Parameter Dan Uji Goodness of Fit Untuk Data Biner Berpasangan," *J. Mat. Komputasi dan Stat.*, vol. 2, no. 1, pp. 1–12, 2022, doi: 10.33772/jmks.v2i1.7.
- [19] S. T. Nihan, "Karl Pearsons chi-square tests," *Educ. Res. Rev.*, vol. 15, no. 9, pp. 575–580, 2020, doi: 10.5897/err2019.3817.
- [20] S. H. Nur, "Model Regresi Cox Proportional Hazard dengan Penaksiran Parameter Efron Partial Likelihood," vol. 6, no. 2, pp. 192–203, 2024, doi: 10.31605/jomta.v6i2.3907.
- [21] I. N. Wardani and A. Sofro, "Estimasi Parameter Model Regresi Probit Multinomial Dengan Metode Maximum Likelihood," *MATHunesa J. Ilm. Mat.*, vol. 8, no. 2, pp. 209–215, 2020, doi: 10.26740/mathunesa.v8n2.p209-215.
- [22] F. Arofah and A. Sofro, "Penerapan Regresi Logistik Multinomial untuk Analisis Model Tingkat Depresi pada Lansia," *MATHunesa J. Ilm. Mat.*, vol. 10, no. 1, pp. 84–93, 2022, doi: 10.26740/mathunesa.v10n1.p84-93.
- [23] A. George, T. S. Stead, and L. Ganti, "What's the Risk: Differentiating Risk Ratios, Odds Ratios, and Hazard Ratios?," *Cureus*, vol. 12, no. 8, pp. 6–13, 2020, doi: 10.7759/cureus.10047.

- [24] R. Ilmi Hidayati, “Analisis Imbalanced Multiclass Pada Status Kepemilikan Asuransi Dengan Metode Multinomial Logistic Regression,” 2019, [Online]. Available: <http://repository.its.ac.id/id/eprint/64492>
- [25] P. A. R. . Putri, S. S. . Prasetyowati, and Y. . Sibaroni, “The Performance of the Equal-Width and Equal-Frequency Discretization Methods on Data Features in Classification Process ”, *Sinkron*, vol. 7, no. 4, pp. 2082-2098, 2023, doi: 10.33395/sinkron.v8i4.12730.