

Received: 30 April 2022

Revised: 15 June 2022

Accepted: 28 June 2022

Published: 30 June 2022

Pemodelan Regresi Logistik Biner pada Faktor-Faktor yang Memengaruhi Indeks Pembangunan Manusia di Pulau Jawa

Ilham Alifa Azagi^{1, a)}, Erfiani^{1, b)}, Indahwati^{1, c)}, Anwar Fitrianto^{1, d)}, Reni Amelia^{1, e)}

¹Prodi Statistika dan Sains Data, IPB University

Email: ^{a)}ilham29alifa@apps.ipb.ac.id, ^{b)}erfiani@apps.ipb.ac.id, ^{c)}indah.stk@gmail.com,
^{d)}anwarstat@gmail.com, ^{e)}reniamelia@apps.ipb.ac.id

Abstract

The Human Development Index is built from three basic dimensions. The dimension as the basis for HDI is an approach chosen in describing the quality of human life and the absence of change. These three dimensions consist of a long and healthy life, knowledge, and a decent standard of living. This study aims to determine a binary logistic regression model and the factors that influence the Human Development Index in Java using binary logistic regression. Binary logistic regression is a statistical analysis used to determine the effect of one or more independent variables on a dependent variable, where the dependent variable is binary. The results of this study indicate that the binary logistic regression model formed is $Y = -111.32025 + 0.58148X_1 + 3.07145X_2 + 2.40689X_3 + 0.79014X_4 - 0.87295X_5 + 0.02139X_6 + 0.11445X_7$. The independent variables that significantly affect the HDI with a significance level of $\alpha = 5\%$ are the percentage of life expectancy (X_1), the percentage of expected years of schooling (X_2), and the percentage of the average length of schooling (X_3). Evaluation of the model obtained a balanced accuracy value of 96.43%, which means that the model is good to use.

Keywords: HDI, binary logistic regression, Java, evaluation model

Abstrak

Indeks Pembangunan Manusia dibangun dari tiga dimensi dasar. Dimensi sebagai dasar IPM adalah suatu pendekatan yang terpilih dalam menggambarkan kualitas hidup manusia dan tidak adanya perubahan. Ketiga dimensi ini terdiri dari umur panjang dan hidup sehat, pengetahuan, dan standar hidup layak. Penelitian ini bertujuan untuk melakukan pemodelan regresi logistik biner dan mengetahui faktor apa saja yang mempengaruhi Indeks Pembangunan Manusia di Pulau Jawa. Regresi logistik biner merupakan suatu metode statistika yang digunakan untuk mengetahui adanya pengaruh satu atau lebih variabel bebas terhadap satu variabel terikat, dimana variabel terikat bersifat biner. Hasil penelitian ini menunjukkan model yang diperoleh adalah $Y = -111.32025 + 0.58148X_1 + 3.07145X_2 + 2.40689X_3 + 0.79014X_4 - 0.87295X_5 + 0.02139X_6 + 0.11445X_7$. Variabel bebas yang mempengaruhi secara signifikan terhadap IPM

taraf nyata $\alpha = 5\%$ adalah Persentase umur harapan hidup (X1), Persentase harapan lama sekolah (X2), dan Persentase rata-rata lama sekolah (X3). Evaluasi model diperoleh nilai balanced accuracy sebesar 96.43%, yang berarti model baik digunakan.

Kata-kata kunci: HDI, regresi logistik biner, Jawa, evaluasi model.

PENDAHULUAN

Salah satu permasalahan yang terjadi di Indonesia yakni Indeks Pembangunan Manusia (IPM). Indeks Pembangunan Manusia merupakan salah satu indikator yang penting dalam melihat sisi lain dari pembangunan. Setiap indikator komponen penghitungan IPM dapat dimanfaatkan untuk mengukur keberhasilan pembangunan kualitas hidup manusia (masyarakat/penduduk) (BPS, 2020). Keberhasilan pembangunan khususnya pembangunan manusia dapat dinilai secara parsial dengan melihat seberapa besar permasalahan yang paling mendasar di masyarakat dapat teratasi. Namun persoalannya adalah capaian pembangunan manusia secara parsial sangat bervariasi dimana beberapa aspek pembangunan tertentu berhasil dan beberapa aspek pembangunan lainnya gagal (Nurmalasari, 2017).

Pembangunan di suatu daerah merupakan upaya untuk meningkatkan kesejahteraan dan mewujudkan kemakmuran masyarakat yang diukur menggunakan indikator IPM. IPM dibangun dari tiga dimensi dasar, dimensi ini merupakan pendekatan yang dipilih dalam penggambaran kualitas hidup manusia dan tidak mengalami perubahan hingga saat ini. Ketiga dimensi ini mencakup umur panjang dan hidup sehat, pengetahuan, dan standar hidup layak (BPS, 2020). Hal ini menjadi tolak ukur untuk meneliti menggunakan dimensi tersebut menjadi variabel yang akan digunakan dalam meneliti IPM di Pulau Jawa.

Pulau Jawa mencakup enam Provinsi, Provinsi tersebut yakni Provinsi DKI Jakarta, Provinsi Banten, Provinsi Jawa Barat, Provinsi Jawa Tengah, Provinsi DI Yogyakarta, dan Provinsi Jawa Timur. IPM di Pulau Jawa pada tahun 2020 sebesar 72,71. Badan Pusat Statistik (BPS) mengategorikan nilai IPM menjadi empat kategori yakni sangat tinggi, tinggi, sedang, dan rendah. IPM pada setiap kabupaten/kota di Pulau Jawa masih belum merata dan terdapat kesenjangan yang cukup tinggi. Nilai IPM tertinggi di Pulau Jawa yaitu 86,61 dan terendah yaitu 62,7 maka akan dilakukan penelitian untuk mengetahui faktor-faktor yang mempengaruhi IPM di Pulau Jawa.

Faktor-faktor di setiap sektor pembentuk IPM akan cenderung memiliki hubungan yang kuat satu sama lainnya, karena antar faktor-faktor tersebut saling mempengaruhi (Putra & Ratnasari, 2015). Faktor tergantung dari variabel bebas dan variabel terikat. Variabel terikat yang digunakan adalah nilai IPM. Nilai IPM dibagi menjadi dua kategori dimana nilai IPM di Pulau Jawa tidak terdapat nilai yang rendah dan untuk nilai IPM tinggi dan sangat tinggi akan menjadi satu kategori. Hal ini membuat penelitian ini akan menggunakan dua kategori.

Penelitian dengan dua kategori yaitu menggunakan Regresi Logistik Biner. Analisis regresi logistik biner adalah analisis statistika yang digunakan untuk mengetahui pengaruh satu atau lebih variabel bebas terhadap satu variabel terikat, dimana variabel terikat bersifat biner atau *dichotomous*. Variabel *dichotomous* adalah variabel yang bersifat kategori dengan jumlah kategori sebanyak dua, misalnya berhasil dan gagal (Fibriyani & Ani, 2014). Tujuan dari penelitian ini: (1) menentukan model regresi logistik Indeks Pembangunan Manusia (IPM) di Pulau Jawa; (2) menentukan faktor yang mempengaruhi Indeks Pembangunan Manusia (IPM) di Pulau Jawa.

METODOLOGI

Bahan dan Data

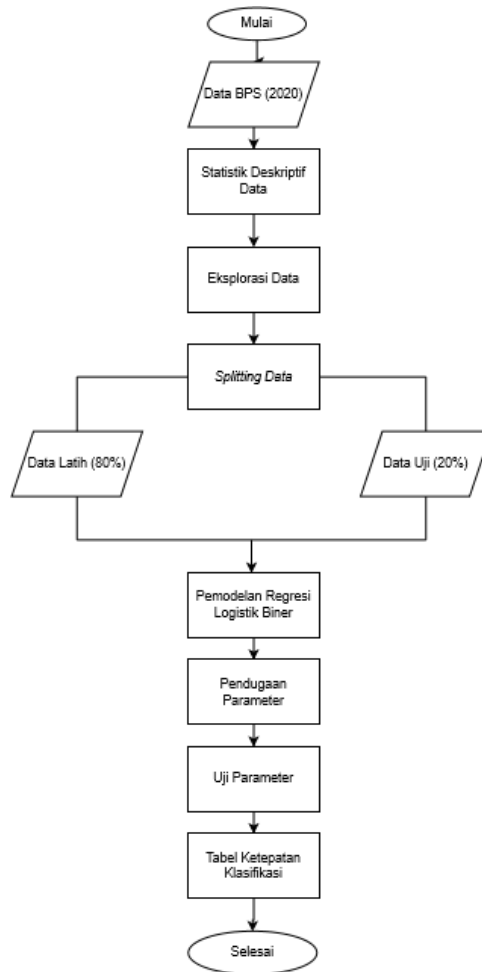
Data yang digunakan adalah data sekunder. Data bersumber dari Badan Pusat Statistik (BPS). Data dari publikasi yang berjumlah 119 data Kabupaten/Kota di Pulau Jawa. Pulau Jawa terdiri dari Provinsi Banten, Provinsi DKI Jakarta, Provinsi Jawa Barat, Provinsi Jawa Tengah, Provinsi DI Yogyakarta, dan Provinsi Jawa Timur.

Variabel yang digunakan dalam penelitian ini adalah variabel terikat dan variabel bebas. Variabel terikat disimbolkan dengan Y dan variabel bebas disimbolkan dengan X . Variabel respons menunjukkan nilai IPM yang telah dikategorikan yaitu tinggi dan sedang. Tabel 1 menjelaskan beberapa literatur yang menjelaskan hubungan beberapa variabel dengan nilai IPM.

TABEL 1. Variabel Penelitian dan Referensi

No.	Variabel Penelitian	Referensi	Kesimpulan
1	Kategori Nilai IPM (Y)		
2	Umur Harapan Hidup (%) (X_1)		Semakin tinggi umur harapan hidup akan meningkatkan kecenderungan untuk nilai IPM
3	Harapan Lama Sekolah (%) (X_2)	BPS (2020)	Semakin tinggi harapan lama sekolah akan meningkatkan kecenderungan untuk nilai IPM
4	Rata-rata Lama Sekolah (%) (X_3)		Semakin tinggi rata-rata lama sekolah akan meningkatkan kecenderungan untuk nilai IPM
5	Kepadatan Penduduk ($\text{km}^2/1000$) (X_4)		Semakin tinggi kepadatan penduduk akan meningkatkan kecenderungan untuk nilai IPM
6	Laju Pertumbuhan Penduduk (%) (X_5)		Semakin tinggi laju pertumbuhan penduduk akan meningkatkan kecenderungan untuk nilai IPM
7	Jumlah Puskesmas Keliling (X_6)	Fibriyani & Ani (2014)	Semakin tinggi jumlah puskesmas keliling akan meningkatkan kecenderungan untuk nilai IPM
8	<i>Sex Ratio</i> (%) (X_7)		Semakin tinggi <i>sex ratio</i> akan meningkatkan kecenderungan untuk nilai IPM

Metode Penelitian



GAMBAR 1. Alur tahapan penelitian

Berdasarkan Gambar 1, tahapan awal yang dilakukan adalah mengumpulkan data, lalu melakukan deskripsi data, eksplorasi data, membagi data menjadi data latih dan data uji, melakukan pemodelan regresi logistik biner, melakukan pendugaan parameter, menguji signifikansi parameter, dan melakukan tabel ketepatan klasifikasi.

Regresi Logistik Biner

Regresi logistik adalah suatu metode analisis statistika untuk mendeskripsikan hubungan antara peubah respons yang memiliki dua atau lebih kategori dengan satu atau lebih peubah penjelas yang berskala kategori atau interval (Hosmer & Lemeshow, 2013). Regresi logistik terdiri dari regresi logistik biner, regresi logistik ordinal, dan regresi logistik multinomial.

Model regresi logistik secara tidak langsung memodelkan variabel bebas berdasarkan probabilitas yang terkait dengan nilai pada variabel terikat. Probabilitas yang digunakan adalah $\pi(x)$ ketika $Y = 1$ dan $1 - \pi(x)$ ketika $Y = 0$. Distribusi log (atau transformasi logistik dari π) disebut juga logit dari π atau logit (π) yang merupakan log (berdasarkan e) dari odds ratio variabel terikatnya adalah satu. Didefinisikan sebagai berikut (Elsalam, 2015):

$$\text{logit}(\pi) = \log\left(\frac{\pi}{1-\pi}\right) = \ln\left(\frac{\pi}{1-\pi}\right)$$

Nilai π hanya berkisar dari 0 hingga 1, skala logit (π) berkisar dari tak hingga negatif hingga tak hingga positif dan simetris di sekitar 0.5. Bentuk persamaan regresi logistik sebagai berikut (Elsalam, 2015):

$$\text{logit } \pi(x) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \dots$$

Fungsi link probit adalah fungsi kumulatif normal terbaik. Log-log komplementer adalah fungsi log natural (Myers *et al*, 2010). Akan tetapi, fungsi logit link paling banyak digunakan karena tidak terlalu rumit dan mudah untuk diinterpretasikan (Faraway, 2016).

Bentuk model regresi logistik dengan satu variabel respons adalah :

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}$$

Pendugaan Parameter

Pendugaan parameter untuk model regresi logistik biner dapat dilakukan dengan metode maksimum likelihood. Metode maksimum likelihood yaitu metode kuadrat terkecil tertimbang yang dilakukan beberapa iterasi. Metode maksimum likelihood akan menghasilkan suatu nilai dari parameter yang tidak diketahui yang memaksimalkan kemungkinan memperoleh kumpulan data teramati. Penduga yang dihasilkan adalah penduga paling mendekati dengan data yang diamati.

Apabila peubah terikat (Y) dikodekan sebagai 0 atau 1 maka pernyataan diberikan untuk persamaan regresi logistik memberikan probabilitas bersyarat bahwa Y sama dengan 1 diberikan x . Ini dilambangkan sebagai $P(Y = 1|x)$, maka kuantitas $1 - \pi(x)$ memberikan probabilitas bersyarat bahwa Y sama dengan nol yang diberi x , $P(Y = 0|x)$. Kemudian, untuk pasangan (x_i, y_i) , dimana $y_i = 1$, kontribusi untuk fungsi likelihood adalah $\pi(x_i)$, dan untuk pasangan itu dimana $y_i = 0$, kontribusi untuk fungsi likelihood adalah $1 - \pi(x_i)$, dimana kuantitas $\pi(x_i)$ menunjukkan nilai $\pi(x)$ dihitung pada x_i (Safitri, 2019).

Ketepatan Klasifikasi

Ada beberapa cara mengevaluasi model regresi logistik, salah satunya dengan melihat performa klasifikasi. Ketepatan klasifikasi dengan melihat nilai pada *Balanced Accuracy*. Evaluasi performa klasifikasi sebagai berikut (Amelia, 2021):

TABEL 2. Ketepatan Klasifikasi

Prediksi	Observasi	
	0	1
0	a	b
1	c	d

$$\text{Balanced Accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2}$$

Dimana:

$$\text{Sensitivity} = a/(a + c)$$

$$\text{Specificity} = d/(d + b)$$

HASIL DAN PEMBAHASAN

Statistika Deskriptif

Nilai IPM Kabupaten/Kota di Pulau Jawa tahun 2020 berada pada nilai 72,71, dimana masuk dalam kategori tinggi. Untuk itu dilakukan pembentukan variabel terikat baru dengan mengategorikan. Pengkategorian dilakukan menurut Badan Pusat Statistik (BPS), yaitu kategori 0 ($Y=0$) untuk Kabupaten/Kota yang memiliki nilai IPM sedang ($60 \leq IPM \leq 70$) di Pulau Jawa, sedangkan kategori 1 ($Y=1$) untuk Kabupaten/Kota yang memiliki nilai IPM tinggi dan sangat tinggi ($IPM \geq 70$) di Pulau Jawa.

TABEL 3. Statistik Deskriptif Variabel Terikat

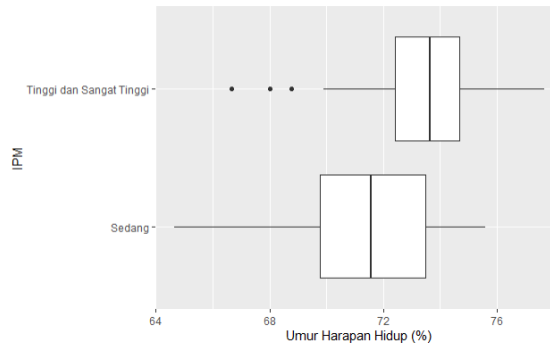
Kategori	N	Persentase
Kategori 0	43	36,13
Kategori 1	76	63,87
Total	119	100

TABEL 4. Statistik Deskriptif Variabel Bebas

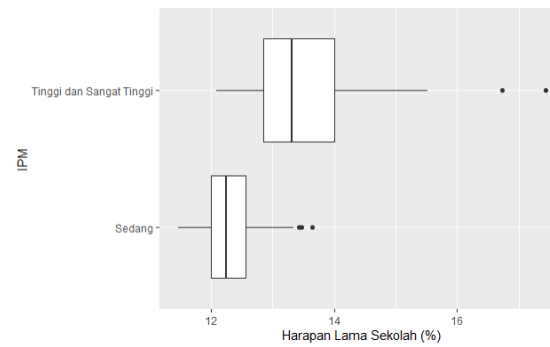
Variabel	N	Mean	Standar Deviasi	Minimum	Maksimum
X_1	119	72,7	2,609	64,64	77,65
X_2	119	13,14	1,034	11,46	17,43
X_3	119	8,166	1,735	1,63	11,81
X_4	119	3,31	4,51	0,279	20,813
X_5	119	0,9278	0,693	0,09	3,95
X_6	119	38,71	48,776	3	379
X_7	119	99,11	3,141	90,73	106,2

Eksplorasi Data Awal

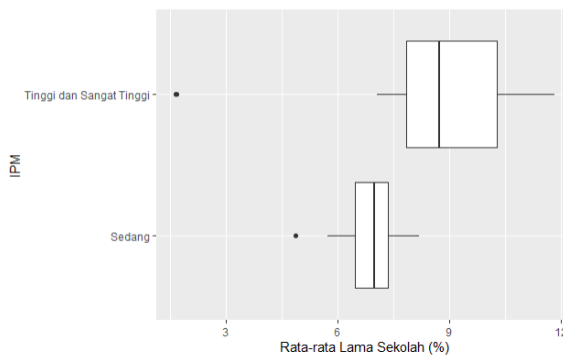
Data sampel memiliki 119 Kabupaten/Kota di Pulau Jawa pada tahun 2020, persentase umur harapan hidup memiliki median nilai IPM tinggi dan sangat tinggi lebih besar dibandingkan median nilai IPM sedang (Gambar 2). Persentase harapan lama sekolah memiliki median nilai IPM tinggi dan sangat tinggi lebih besar dibandingkan median nilai IPM sedang (Gambar 3). Persentase rata-rata lama sekolah memiliki median nilai IPM tinggi dan sangat tinggi lebih besar dibandingkan median nilai IPM sedang (Gambar 4). Kepadatan penduduk memiliki median nilai IPM tinggi dan sangat tinggi lebih besar dibandingkan median nilai IPM sedang (Gambar 5). Persentase laju pertumbuhan penduduk memiliki median nilai IPM tinggi dan sangat tinggi lebih besar dibandingkan median nilai IPM sedang (Gambar 6). Jumlah puskesmas keliling memiliki median nilai IPM tinggi dan sangat tinggi lebih kecil dibandingkan median nilai IPM sedang (Gambar 7). Persentase *sex ratio* memiliki median nilai IPM tinggi dan sangat tinggi lebih besar dibandingkan median nilai IPM sedang (Gambar 8).



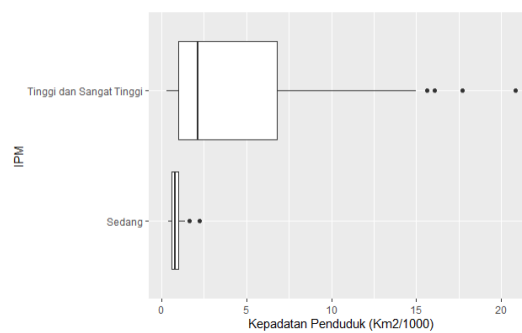
GAMBAR 2. Persentase Umur Harapan Hidup



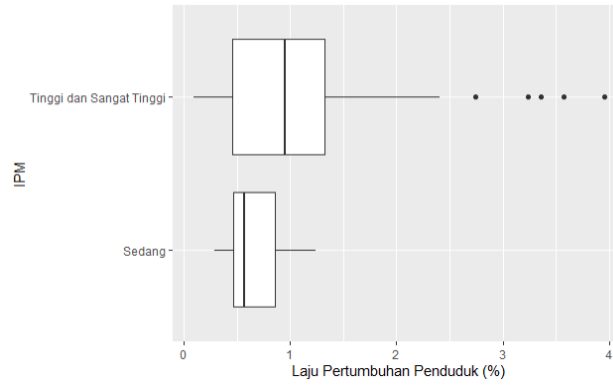
GAMBAR 3. Persentase Harapan Lama Sekolah



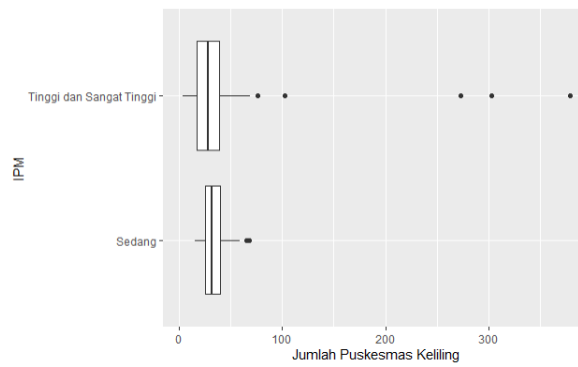
GAMBAR 4. Persentase Rata-rata Lama Sekolah



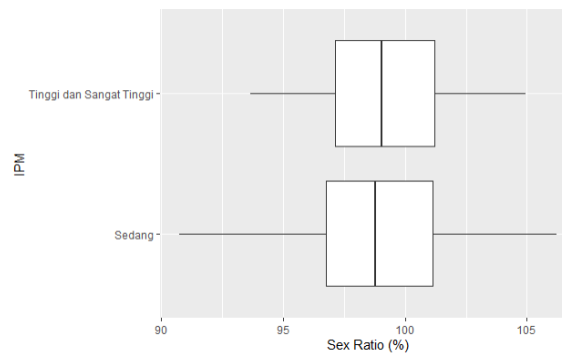
GAMBAR 5. Kepadatan Penduduk



GAMBAR 6. Persentase Laju Pertumbuhan Penduduk



GAMBAR 7. Jumlah Puskesmas Keliling



GAMBAR 8. Persentase *Sex Ratio*

Membagi Data Menjadi 80% Data Latih dan 20% Data Uji

Membagi data dengan memilih 80% dari data latih dan 20% dari data uji. Terdapat 95 Kabupaten/Kota yang menjadi data latih (*training*) dan 24 Kabupaten/Kota yang menjadi data uji (*testing*).

Pemodelan Regresi Logistik Biner

Sebelum memodelkan regresi logistik biner dilakukan dengan melihat korelasi atau hubungan antar variabel bebas. Hal ini dilakukan untuk melihat ada atau tidaknya kolinearitas antar variabel bebas. Pengujian ini dimaksudkan agar tidak terjadi pelanggaran asumsi yaitu tidak terdapat multikolinieritas dalam pemodelan regresi logistik biner. Berdasarkan Tabel 5 diperoleh bahwa tidak ada koefisien korelasi yang lebih besar dari 0.95, sehingga dapat dikatakan bahwa tidak terdapat kolinearitas antar variabel bebas.

TABEL 5. Korelasi Antar Variabel Bebas

	X_2	X_3	X_4	X_5	X_6	X_7
X_1	0.316	0.309	0.244	0.048	-0.182	-0.219
X_2		0.590	0.414	0.176	-0.017	-0.193
X_3			0.458	0.350	0.074	0.081
X_4				0.355	0.216	0.139
X_5					0.503	0.376
X_6						0.357

Selanjutnya melakukan pemodelan regresi logistik biner. Pemodelan ini dilakukan dengan menggunakan data latih. Untuk data uji digunakan saat melihat ketepatan klasifikasi pada data. Tabel 6 diperoleh hasil regresi logistik biner.

TABEL 6. Hasil Regresi Logistik Biner

Variabel	Estimasi Koefisien	Standar Error	Statistik Uji Z	Nilai Probabilitas	Odds Ratio
X_1	0.58148	0.26504	2.194	0.02824*	1.789
X_2	3.07415	1.24792	2.463	0.01376*	21.631
X_3	2.40689	1.06143	2.268	0.02335*	11.099
X_4	0.79014	0.55860	1.414	0.15722	2.204
X_5	-0.87295	1.37717	-0.634	0.52616	0.418
X_6	0.02139	0.04056	0.527	0.59798	1.022
X_7	0.11445	0.19869	0.576	0.56462	1.121

Pengujian Parameter

Pengujian parameter dengan melihat variabel bebas yang berpengaruh signifikan terhadap nilai IPM, dimana hipotesis yang akan diuji adalah:

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0, i = 1, 2, \dots, 7$$

Berdasarkan Tabel 6 diperoleh bahwa variabel bebas yang berpengaruh signifikan pada taraf nyata $\alpha = 5\%$ terhadap nilai IPM di Kabupaten/Kota di Pulau Jawa adalah Persentase umur harapan hidup (X_1), Persentase harapan lama sekolah (X_2), dan Persentase rata-rata lama sekolah (X_3).

Sehingga model regresi logistik biner yang terbentuk adalah

$$Y = -111.32025 + 0.58148X_1 + 3.07145X_2 + 2.40689X_3 + 0.79014X_4 - 0.87295X_5 + 0.02139X_6 + 0.11445X_7$$

Interpretasi dari odds ratio yaitu:

- $Exp(\text{Persentase umur harapan hidup}) = 1.789$ berarti setiap peningkatan 1% umur harapan hidup di suatu Kabupaten/Kota menyebabkan Kabupaten/Kota tersebut memiliki kecenderungan untuk memperoleh nilai IPM tinggi dan sangat tinggi sebesar 1.789 kali dibandingkan dengan nilai IPM sedang.
- $Exp(\text{Persentase harapan lama sekolah}) = 21.631$ berarti setiap peningkatan 1% harapan lama sekolah di suatu Kabupaten/Kota menyebabkan Kabupaten/Kota tersebut memiliki kecenderungan untuk memperoleh nilai IPM tinggi dan sangat tinggi sebesar 21.631 kali dibandingkan dengan nilai IPM sedang.
- $Exp(\text{Persentase rata-rata lama sekolah}) = 11.099$ berarti setiap peningkatan 1% rata-rata lama sekolah di suatu Kabupaten/Kota menyebabkan Kabupaten/Kota tersebut memiliki kecenderungan untuk memperoleh nilai IPM tinggi dan sangat tinggi sebesar 11.099 kali dibandingkan dengan nilai IPM sedang.
- $Exp(\text{Kepadatan penduduk}) = 2.204$ berarti setiap peningkatan 1 Km² kepadatan penduduk di suatu Kabupaten/Kota menyebabkan Kabupaten/Kota tersebut memiliki kecenderungan untuk memperoleh nilai IPM tinggi dan sangat tinggi sebesar 2.204 kali dibandingkan dengan nilai IPM sedang.
- $Exp(\text{Persentase laju pertumbuhan penduduk}) = 0.418$ berarti setiap peningkatan 1% laju pertumbuhan penduduk di suatu Kabupaten/Kota menyebabkan Kabupaten/Kota tersebut memiliki kecenderungan untuk memperoleh nilai IPM tinggi dan sangat tinggi sebesar 0.418 kali dibandingkan dengan nilai IPM sedang.
- $Exp(\text{Jumlah puskesmas keliling}) = 1.022$ berarti setiap peningkatan 1 puskesmas keliling di suatu Kabupaten/Kota menyebabkan Kabupaten/Kota tersebut memiliki kecenderungan untuk memperoleh nilai IPM tinggi dan sangat tinggi sebesar 1.022 kali dibandingkan dengan nilai IPM sedang.
- $Exp(\text{Persentase sex ratio}) = 1.121$ berarti setiap peningkatan 1% *sex ratio* di suatu Kabupaten/Kota menyebabkan Kabupaten/Kota tersebut memiliki kecenderungan untuk memperoleh nilai IPM tinggi dan sangat tinggi sebesar 1.121 kali dibandingkan dengan nilai IPM sedang.

Evaluasi Model dengan Melihat Tabel Ketepatan Klasifikasi

Evaluasi model dengan melihat nilai pada *balanced accuracy*. Ketepatan klasifikasi dilihat dari data uji pada model. Tabel 7 diperoleh nilai *balanced accuracy* lebih dari 80%. Hal ini berarti model baik digunakan untuk melakukan pemodelan regresi logistik biner.

TABEL 7. Hasil Ketepatan Klasifikasi

Ketepatan	Nilai	Persentase
<i>Accuracy</i>	0.9583	95.83%
<i>Sensitivity</i>	0.9286	92.86%
<i>Specificity</i>	1	100%
<i>Balanced Accuracy</i>	0.9643	96.43%

KESIMPULAN DAN SARAN

Hasil penelitian ini menunjukkan model regresi logistik biner yang terbentuk adalah $Y = -111.32025 + 0.58148X_1 + 3.07145X_2 + 2.40689X_3 + 0.79014X_4 - 0.87295X_5 + 0.02139X_6 +$

$0.11445X_7$. Faktor yang berpengaruh terhadap kategori pada nilai IPM terhadap variabel bebas dengan taraf nyata $\alpha = 5\%$ adalah Persentase umur harapan hidup (X_1), Persentase harapan lama sekolah (X_2), dan Persentase rata-rata lama sekolah (X_3). Evaluasi model diperoleh nilai *balanced accuracy* sebesar 96.43%, yang berarti model baik digunakan dalam pembentukan model regresi logistik biner. Untuk penelitian selanjutnya dapat digunakan penambahan faktor spasial atau wilayah, dikarenakan kondisi pada Pulau Jawa. Analisis yang akan digunakan yaitu *Geographically Weighted Logistic Regression* (GWLR).

REFERENSI

- Amelia, Reni. (2021). Komparasi Teknik Undersampling dan Oversampling Pada Regresi Logistik Biner Dalam Menduga Faktor Determinan Berhenti Merokok Penduduk Lanjut Usia. *Jurnal TIMES*. 10(2) : 1-11.
- Badan Pusat Statistik (BPS). (2020). Indeks Pembangunan Manusia. Jakarta : Badan Pusat Statistik.
- Elsalam, N.M.M.A. (2015). Binary Logistic Regression to Identify the Risk Factors of Eye Glaucoma. *International Journal of Sciences: Basic and Applied Research*. 23(1) : 366-376.
- Faraway, J.J. (2016). *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. USA : CRC Press.
- Fibriyani, V., & Ani, A. (2014). Pemodelan Indeks Pembangunan Manusia Kabupaten/Kota di Provinsi Papua dan Papua Barat Tahun 2014. *Jurnal Varian*. 1(2) : 30-40.
- Hosmer, D. W., & Lemeshow, S. (2013). *Applied Logistic Regression (Third Edition)*. New York : John. Wiley & Sond, INC.
- Myers, R.H., Montgomery, D.C., & Vining, G.G. (2010). *Generalized Linear Models: With Applications in Engineering and the Sciences*. New York : Wiley.
- Nurmalasari, R. (2017). Analisis Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia (IPM) Menggunakan Metode Regresi Logistik Ordinal dan Regresi Probit Ordinal (Studi Kasus Kabupaten/Kota di Jawa Tengah Tahun 2014). *Jurnal Gaussian*. 6(1) : 111-120.
- Putra, D. M., & Vita, R. (2015). Pemodelan Indeks Pembangunan Manusia (IPM) Provinsi Jawa Timur dengan Menggunakan Metode Regresi Logistik Ridge. *Jurnal Sains dan Seni ITS*. 4(2) : 175-180.
- Safitri, A. (2019). Model Regresi Logistik Biner pada Tingkat Pengangguran Terbukadi Provinsi Sulawesi Barat Tahun 2017. *Journal of Statistics and Its Application on Teaching and Research*. 1(2) : 1-6.