

## PEMODELAN INDEKS PEMBANGUNAN MANUSIA PROVINSI JAWA BARAT, JAWA TIMUR DAN JAWA TENGAH TAHUN 2019 DENGAN MENGGUNAKAN METODE REGRESI LOGISTIK ORDINAL

Meylita Sari<sup>1</sup> dan Purhadi<sup>2</sup>

<sup>1,2</sup>Departemen Statistika, Institut Teknologi Sepuluh Nopember  
Kampus ITS, Sukolilo, Surabaya 60111  
\*email: [meylitasari27@gmail.com](mailto:meylitasari27@gmail.com)

### ABSTRACT

*Ordinal logistic regression is one of the statistical methods to analyze response variables (dependents) that have an ordinal scale consisting of three or more categories. Predictor variables (independent) that can be included in the model are category or continuous data consisting of two or more variables. Human Development Index (HDI) is an indicator of the success of human development in a region and can be categorized into medium, high and very high. Based on the further categorization, in this study would like to know more about the HDI model using the Ordinal Logistic Regression method, with predictor variables that are suspected to affect, so that it is obtained in West Java Province is influenced by variable poverty rates and clean water sources with a classification accuracy value of 77.78%, Central Java Province is influenced by variable economic growth rate based on constant price GDP, poverty rate and open unemployment rate with a classification accuracy value of 82.85%. East Java province is influenced by variable poverty rate and open unemployment rate with a classification accuracy value of 76.31%. As well as in the three provinces in Java Island is influenced by variable economic growth rate, variable poverty rate, variable clean water source with a classification accuracy value of 73%.*

**Keywords :** *Ordinal Logistic Regression, HDI, Classification Accuracy*

### 1. PENDAHULUAN

Posisi manusia selalu menjadi tema sentral dalam setiap program pencapaian pembangunan. Dalam skala internasional dikenal tujuan pembangunan milenium (Millenium Development Goals/MDG's), yang disepakati oleh pemimpin dunia dalam KTT 2000. MDG merupakan komitmen masyarakat internasional, khususnya negara yang sedang berkembang, terhadap visi pembangunan. Visi ini secara kuat menempatkan pembangunan sosial dan ekonomi secara berkelanjutan. Sementara paradigma pembangunan Indonesia mengalami perkembangan sejak tahun 1970 sampai dengan saat ini. Oleh karena itu, selama kurun waktu 1970-1980, pemerintah berupaya keras untuk menerapkan paradigma pemerataan pembangunan di seluruh wilayah. Pemerintah secara sentralistik menerapkan program-program pembangunan kepada daerah-daerah miskin dan pelosok-pelosok desa untuk mengejar ketertinggalan. Hasil akhir dari penelitian yang dilakukan oleh Zairin (2006) yaitu pembangunan manusia tersebut adalah lahirnya manusia yang mandiri dan mampu memberikan kontribusi terhadap keberlanjutan pembangunan nasional di seluruh wilayah. Sejalan dengan perkembangan pendekatan pembangunan manusia di atas, pada tahun 1990 PBB melalui UNDP menetapkan sebuah tolok ukur untuk mengukur hasil-hasil pembangunan manusia dengan indikator IPM (Indeks Pembangunan Manusia). Keberhasilan pembangunan khususnya pembangunan manusia dapat dinilai secara parsial dengan melihat seberapa besar permasalahan yang paling mendasar di masyarakat dapat teratasi. Permasalahan-permasalahan tersebut diantaranya adalah kemiskinan, pengangguran, buta huruf, ketahanan pangan, dan penegakan demokrasi (BPS, 2013).

Penelitian tentang IPM telah banyak dilakukan sebelumnya, diantaranya adalah penelitian Citra (2010) yang meneliti kesamaan vektor parameter pada beberapa model

regresi logistik ordinal dengan studi kasus faktor-faktor yang mempengaruhi IPM di provinsi Jawa Timur, Jawa Tengah, Jawa Barat dan Sumatera Utara pada tahun 2007. Data IPM yang digunakan merupakan data kualitatif yang dikategorikan menjadi 3 kategori dengan faktor-faktor yang diduga mempengaruhinya berupa data kualitatif maupun kuantitatif. Karena pada penelitian ini variabel respon berbentuk kategori bertingkat atau ordinal terdapat metode yang sesuai yaitu metode Regresi Logistik Ordinal.

Penelitian ini dilakukan untuk mengetahui faktor yang akan dianalisis diantaranya Laju Pertumbuhan PDRB berdasarkan Harga Konstan, Persentase Penduduk Miskin, Persentase Rumah Tangga Dengan Akses Sumber Air Minum Layak, Tingkat Pengangguran Terbuka (TPT), dan Tingkat Partisipasi Angkatan Kerja (TPAK) yang mempengaruhi IPM dari Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur pada tahun 2019 dengan menggunakan metode Regresi Logistik Ordinal.

## 2. TINJAUAN PUSTAKA

### 2.1 Regresi Logistik Ordinal

Regresi logistik ordinal merupakan salah satu metode statistika untuk menganalisis variabel respon (dependen) yang mempunyai skala ordinal yang terdiri atas tiga kategori atau lebih. Variabel prediktor (independen) yang dapat disertakan dalam model berupa data kategori atau kontinu yang terdiri atas dua variabel atau lebih. Model regresi logistik ordinal dibentuk dari distribusi multinomial dengan  $n=1$ , diberikan probabilitas dari distribusi Multinomial untuk  $n=1$  sebagai berikut:

$$P(Y_1 = y_1, Y_2 = y_2) = p_1^{y_1} p_2^{y_2} (1 - p_1 - p_2)^{1-y_1-y_2}, \quad (Y_1, Y_2) \sim M(1, p_1, p_2)$$

$$0 < p_1 < 1; 0 < p_2 < 1; 0 \leq y_1 + y_2 \leq 1$$

Kemudian dengan menggunakan probabilitas dari distribusi multinomial didapatkan model logit. Model logit yang menjadi fokus utama dalam bidang sosial adalah model logit kumulatif (*cumulative logit models*). Jika variabel prediktor  $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_p]^T$ , maka peluang kumulatif logit didefinisikan (Agresti, 2002) sebagai

$$P(Y \leq j | \mathbf{x}) = \pi_1(\mathbf{x}) + \pi_2(\mathbf{x}) + \dots + \pi_j(\mathbf{x}), \quad j = 1, 2, \dots, J$$

Kumulatif logit didefinisikan sebagai

$$\begin{aligned} \text{logit} [P(Y \leq j | \mathbf{x})] &= \ln \left[ \frac{P(Y \leq j | \mathbf{x})}{1 - P(Y \leq j | \mathbf{x})} \right] \\ &= \ln \left[ \frac{P(Y \leq j | \mathbf{x})}{P(Y > j | \mathbf{x})} \right] \\ &= \ln \left[ \frac{\pi_1(\mathbf{x}) + \pi_2(\mathbf{x}) + \dots + \pi_j(\mathbf{x})}{\pi_{j+1}(\mathbf{x}) + \pi_{j+2}(\mathbf{x}) + \dots + \pi_J(\mathbf{x})} \right], j = 1, 2, \dots, J - 1 \end{aligned} \quad (1)$$

Anggap suatu peubah respon multinomial  $Y$  dengan keluaran kategori yang dinyatakan oleh  $1, 2, \dots, J$  dan misalkan  $\mathbf{x}$  menyatakan suatu vektor kovariat berdimensi  $p$ . Dependensi peluang kumulatif  $Y$  terhadap  $\mathbf{x}$  untuk model *proportional odds* sering dinyatakan dalam bentuk

$$\ln \left[ \frac{P(Y \leq j | \mathbf{x})}{P(Y > j | \mathbf{x})} \right] = \beta_{0j} + \boldsymbol{\beta}^T \mathbf{x}, \quad j = 1, 2, \dots, J - 1 \quad (2)$$

Persamaan (2.1) dapat diubah ke dalam bentuk

$$\ln \left[ \frac{\pi_1(\mathbf{x}) + \pi_2(\mathbf{x}) + \dots + \pi_j(\mathbf{x})}{\pi_{j+1}(\mathbf{x}) + \pi_{j+2}(\mathbf{x}) + \dots + \pi_J(\mathbf{x})} \right] = \beta_{0j} + \boldsymbol{\beta}^T \mathbf{x}, \quad j = 1, 2, \dots, J - 1$$

Model yang secara simultan menggunakan semua kumulatif logit (Agresti, 2002) adalah

$$\text{logit} [P(Y \leq j | \mathbf{x})] = \beta_{0j} + \boldsymbol{\beta}^T \mathbf{x}, \quad j = 1, 2, \dots, J - 1$$

$\gamma_j(\mathbf{x}) = P(Y \leq j|\mathbf{x})$  merupakan peluang kumulatif dari kejadian  $(Y \leq j) \cdot \{\beta_{0j}\}$  merupakan parameter intersep yang tidak diketahui yang memenuhi kondisi  $\beta_{01} \leq \beta_{02} \leq \dots \leq \beta_{0,J-1}$  dan  $\boldsymbol{\beta} = [\beta_1 \beta_2 \dots \beta_p]^T$  merupakan vektor koefisien regresi yang tidak diketahui yang bersesuaian dengan  $\mathbf{x}$ .

Jika  $\gamma_j(\mathbf{x}) = \pi_1(\mathbf{x}) + \pi_2(\mathbf{x}) + \pi_3(\mathbf{x}) + \dots + \pi_j(\mathbf{x})$ . Maka  $\gamma_1(\mathbf{x}) = \pi_1(\mathbf{x})$ ,  $\gamma_2(\mathbf{x}) = \pi_1(\mathbf{x}) + \pi_2(\mathbf{x})$ , dan  $\gamma_J(\mathbf{x}) = \pi_1(\mathbf{x}) + \pi_2(\mathbf{x}) + \dots + \pi_J(\mathbf{x}) = 1$ . Model regresi logistik ordinal yang terbentuk jika terdapat  $J$  kategori respon adalah

$$\text{logit } \gamma_1(\mathbf{x}) = \ln \left( \frac{\gamma_1(\mathbf{x})}{1 - \gamma_1(\mathbf{x})} \right) = \beta_{01} + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

$$\text{logit } \gamma_2(\mathbf{x}) = \ln \left( \frac{\gamma_2(\mathbf{x})}{1 - \gamma_2(\mathbf{x})} \right) = \beta_{02} + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

⋮

$$\text{logit } \gamma_{J-1}(\mathbf{x}) = \ln \left( \frac{\gamma_{J-1}(\mathbf{x})}{1 - \gamma_{J-1}(\mathbf{x})} \right) = \beta_{0,J-1} + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

di mana  $\gamma_j(\mathbf{x}) = P(Y \leq j|\mathbf{x}) = \frac{\exp(\beta_{0j} + \boldsymbol{\beta}^T \mathbf{x})}{1 + \exp(\beta_{0j} + \boldsymbol{\beta}^T \mathbf{x})}$ ,  $j = 1, 2, \dots, J - 1$  dan  $\gamma_J(\mathbf{x}) = 1$

Model ini disebut model logistik kumulatif karena rasio odds dari suatu kejadian  $(Y \leq j)$  adalah independen pada setiap indikator kategori.

Jika dimisalkan terdapat 3 kategori respon, maka model regresi logistik ordinal yang terbentuk adalah

$$\text{logit } \gamma_1(\mathbf{x}) = \ln \left( \frac{\gamma_1(\mathbf{x})}{1 - \gamma_1(\mathbf{x})} \right) = \beta_{01} + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

$$\text{logit } \gamma_2(\mathbf{x}) = \ln \left( \frac{\gamma_2(\mathbf{x})}{1 - \gamma_2(\mathbf{x})} \right) = \beta_{02} + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

Selanjutnya untuk mendapatkan pendugaan parameter dapat diperoleh dengan menggunakan metode pendugaan *maximum likelihood estimation (MLE)*. Pendugaan parameter regresi logistik ordinal didapatkan dengan menurunkan fungsi log likelihood terhadap parameter yang akan diestimasi dan disamakan dengan nol. Persamaan  $\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_k}$  digunakan untuk menaksir intersep parameter  $\beta_k$  dimana  $k = 1, 2, \dots, p$  dan  $\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_{\theta_j}}$  digunakan untuk menaksir intersep  $\theta_j$  dimana  $j = 1, 2, \dots, J - 1$ .

Hasil dari persamaan  $\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_k}$  dan  $\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_{\theta_j}}$  merupakan fungsi nonlinear sehingga diperlukan metode iterasi untuk memperoleh estimasi parameternya. Metode iterasi yang dipergunakan yaitu algoritma Newton-Raphson. Iterasi akan berhenti jika terpenuhi kondisi konvergen, yaitu selisih  $\|\boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^{(t)}\| \leq \varepsilon$ , dimana  $\boldsymbol{\theta} = [\beta_1 \beta_2 \dots \beta_p]^T$ ;  $\varepsilon$  adalah bilangan yang sangat kecil.

### Pengujian Parameter

Pengujian parameter dalam regresi logistik ordinal dapat dilakukan baik secara serentak maupun parsial. Uji serentak dilakukan untuk memeriksa peran koefisien  $\boldsymbol{\beta}$  secara keseluruhan atau serentak. Dalam pengujian serentak, uji signifikansi model yang digunakan adalah uji  $G^2$  atau *likelihood-ratio test*.

Hipotesis:

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

$$H_1: \text{Paling sedikit ada satu } \beta_k \neq 0, k=1, 2, \dots, p$$

Statistik uji:

$$G^2 = 2 [\ln L(\hat{\Omega}) - \ln L(\hat{\omega})]$$

di mana  $L(\hat{\omega})$  = nilai yang dimaksimalkan dari fungsi *likelihood* di bawah  $H_0$

$L(\hat{\Omega})$  = nilai yang dimaksimalkan dari fungsi *likelihood* di bawah populasi

Daerah penolakan:

$H_0$  ditolak, jika  $G^2 > \chi^2_{(\alpha, w)}$  di mana  $w$  menunjukkan banyaknya variabel prediktor dalam model atau  $p$ -value  $< \alpha$ .

Untuk pengujian individu yang dapat digunakan adalah *Wald test* (Lea, 1997). Hasil dari *Wald test* ini akan menunjukkan apakah suatu variabel prediktor signifikan atau layak untuk masuk dalam model atau tidak.

Hipotesis:

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0, k=1, 2, \dots, p$$

Statistik uji yang digunakan adalah statistik Wald

$$W_k^2 = \left( \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \right)^2$$

di mana  $\hat{\beta}_k$  merupakan penaksir parameter  $\beta_k$  dan  $SE(\hat{\beta}_k) = \sqrt{\widehat{var}(\hat{\beta}_k)}$ .

Daerah penolakan:

$$H_0 \text{ ditolak jika } W_k^2 > \chi^2_{(\alpha, 1)} \text{ atau } p\text{-value} < \alpha.$$

### Uji Kesesuaian Model (*Goodness of Fit Test*)

Uji ini digunakan untuk menilai apakah model dengan satu atau lebih variabel prediktor merupakan model yang sesuai atau tidak. Statistik uji yang digunakan adalah *deviance*, dengan hipotesis yang digunakan adalah

Hipotesis :

$$H_0: \mu_i = 0$$

(Model regresi logistik ordinal tidak sesuai atau tidak layak digunakan pada data IPM)

$$H_1: \mu_i \neq 0$$

(Model regresi logistik ordinal sesuai atau layak digunakan pada data IPM)

Statistik uji :

$$D = -2 \sum_{i=1}^n \left[ y_{ij} \ln \left( \frac{\hat{\pi}_{ij}}{y_{ij}} \right) + (1 - y_{ij}) \ln \left( \frac{1 - \hat{\pi}_{ij}}{1 - y_{ij}} \right) \right]$$

di mana  $\hat{\pi}_{ij} = \hat{\pi}_j(x_i)$  merupakan peluang observasi ke- $i$  pada kategori ke- $j$

Ringkasan statistik berdasarkan *deviance* mengindikasikan sebaik apa model fit terhadap data. Semakin tinggi nilai  $D$  dan semakin rendah  $p$ -value mengindikasikan bahwa mungkin model tidak fit terhadap data. Derajat bebas untuk uji ini adalah  $J - (p + 1)$  di mana  $J$  merupakan jumlah kovariat dan  $p$  merupakan jumlah variabel prediktor. Jika model adalah terbaik, maka *deviance* akan mendekati distribusi  $\chi^2_{(df)}$ .

### Ketepatan Klasifikasi

Ketepatan klasifikasi adalah suatu evaluasi yang melihat peluang kesalahan yang dilakukan oleh suatu fungsi klasifikasi. Nilai *APER* (*Apparent Error Rate*) menyatakan nilai proporsi sampel yang salah diklasifikasikan oleh fungsi klasifikasi (Johnson dan Wichern, 1992). Penentuan ketepatan pengklasifikasian dapat diketahui melalui tabel klasifikasi sebagai berikut.

$$APER (\%) = \frac{n_{12} + n_{13} + n_{21} + n_{23} + n_{31} + n_{32}}{n} \times 100\%$$

Keterangan:

$n_{11}$  = Jumlah  $Y_i$  dari  $y_i = 1$  tepat diklasifikasikan sebagai  $y_j = 1; i, j = 1, 2, 3$

Ketepatan klasifikasi =  $1 - APER$

## 2.2 Indeks Pembangunan Manusia

Indeks pembangunan manusia merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk). IPM dapat menentukan peringkat atau level pembangunan suatu wilayah/negara. Bagi Indonesia, IPM merupakan data strategis karena selain sebagai ukuran kinerja Pemerintah, IPM juga digunakan sebagai salah satu alokator penentuan Dana Alokasi Umum (DAU). (BPS, 2011).

$$IPM = \sqrt[3]{I_{kesehatan} \times I_{pendidikan} \times I_{pengeluaran}} \times 100$$

Untuk melihat capaian IPM antar wilayah dapat dilihat melalui pengelompokan IPM ke dalam beberapa kategori, yaitu:

IPM < 60 : IPM rendah

60 ≤ IPM < 70 : IPM sedang

70 ≤ IPM < 80 : IPM tinggi

IPM ≥ 80 : IPM sangat tinggi

Definisi untuk variabel independen yang diduga memiliki keterkaitan terhadap perkembangan Indeks Pembangunan Manusia dapat diuraikan sebagai berikut :

1. Laju Pertumbuhan PDRB berdasarkan Harga Konstan,  
Menurut BPS, Salah satu indikator penting untuk mengetahui kondisi ekonomi di suatu negara dalam suatu periode tertentu adalah data Produk Domestik Bruto (PDB) yang dapat digunakan untuk menunjukkan laju pertumbuhan ekonomi secara keseluruhan atau setiap sektor dari tahun ke tahun.
2. Persentase penduduk miskin  
Untuk mengukur kemiskinan, BPS menggunakan konsep kemampuan memenuhi kebutuhan dasar (basic needs approach). Kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Jadi Persentase Penduduk Miskin adalah persentase penduduk yang memiliki rata-rata pengeluaran perkapita perbulan dibawah garis kemiskinan.
3. Persentase rumah tangga dengan akses sumber air bersih  
Menurut BPS Provinsi Jateng (2013) rumah tangga dengan akses air bersih merupakan rumah tangga yang memiliki akses air minum yang bersumber dari air kemasan bermerk, air isi ulang, air leding, sumur bor/pompa, sumur terlindung dan mata air terlindung harus memiliki jarak ≥ 10 meter dari penampungan akhir tinja terdekat.
4. Tingkat Pengangguran Terbuka (TPT)  
Tingkat pengangguran terbuka diukur sebagai persentase jumlah penganggur/pencari kerja terhadap jumlah angkatan kerja.
5. Tingkat Partisipasi Angkatan Kerja (TPAK)  
Menurut BPS Provinsi Jatim (2015) tingkat partisipasi angkatan kerja adalah persentase banyaknya angkatan kerja terhadap penduduk usia kerja.

## 3. METODE PENELITIAN

Sumber data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik masing-masing provinsi di Pulau Jawa yaitu data hasil Survei Sosial Ekonomi Nasional (SUSENAS) Tahun 2019. Variabel yang digunakan dalam penelitian ini terdiri dari variabel dependen (respon) dan variabel independen (prediktor). Variabel dependen (Y) yang dianalisis adalah Indeks Pembangunan Manusia (IPM) pada 119 Kabupaten/Kota dari 6 Provinsi di Pulau Jawa dan menggunakan 5 variabel bebas yang

diduga mempengaruhi IPM. Data IPM dikategorikan menjadi 3 kategori, yaitu 1 = sedang ( $60 \leq \text{IPM} < 70$ ), 2 = tinggi ( $70 \leq \text{IPM} < 80$ ), dan 3 = sangat tinggi ( $\text{IPM} \geq 80$ ). Variabel prediktor yang akan digunakan pada penelitian ini adalah  $X_1$  = Laju Pertumbuhan PDRB berdasarkan Harga Konstan,  $X_2$  = Persentase Penduduk Miskin,  $X_3$  = Persentase Rumah Tangga Dengan Akses Sumber Air Minum Layak,  $X_4$  = Tingkat Pengangguran Terbuka (TPT), dan  $X_5$  = Tingkat Partisipasi Angkatan Kerja (TPAK).

Metode penelitian yang digunakan sebagai langkah-langkah untuk mencapai tujuan penelitian dijabarkan sebagai berikut.

1. Mengumpulkan data yang diperoleh dari Badan Pusat Statistik (BPS) yang digunakan dalam penelitian. Kemudian mengkategorikan variabel yaitu 1 = sedang ( $60 \leq \text{IPM} < 70$ ), 2 = tinggi ( $70 \leq \text{IPM} < 80$ ), dan 3 = sangat tinggi ( $\text{IPM} \geq 80$ ).
2. Melakukan deskriptif data sebagai gambaran pada data IPM Kabupaten/Kota di Pulau Jawa pada tahun 2019 serta faktor-faktor yang diduga memiliki pengaruh terhadap IPM.
3. Uji Multikolinearitas antar variable predictor.
4. Mengetahui pengaruh faktor-faktor Pendidikan, ekonomi dan Kesehatan terhadap IPM di Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur dengan menggunakan regresi logistik ordinal dengan uraian langkah.

Memodelkan variabel  $X_1, X_2, X_3, X_4, X_5$  terhadap Y (variabel tak bebas) dengan metode Regresi Logistik Ordinal.

- a. Menaksir parameter, melakukan pengujian hipotesis serentak dan parsial.
- b. Melakukan pengujian kesesuaian model dengan hipotesis untuk menentukan apakah model regresi logistik ordinal telah sesuai atau tidak.
- c. Interpretasi model untuk faktor-faktor yang berpengaruh terhadap IPM.
- d. Mengetahui hasil ketepatan klasifikasi berdasarkan model regresi Logistik ordinal
  - a. Membangun klasifikasi fungsi peluang dari tiap kategori
  - b. Menghitung presentase klasifikasi dari tiap kategori
  - c. Interpretasi hasil ketepatan klasifikasi

#### 4. HASIL DAN PEMBAHASAN

Nilai IPM yang dikategorikan menjadi 4 kategori seperti pada metodologi tidak terpenuhi, sehingga dilakukan pengkategorian kembali menjadi menjadi 3 kategori, yaitu 1 = sedang ( $60 \leq \text{IPM} < 70$ ), 2 = tinggi ( $70 \leq \text{IPM} < 80$ ), dan 3 = sangat tinggi ( $\text{IPM} \geq 80$ ). Setelah pengkategorian kembali tersebut, seluruh nilai IPM telah terdistribusi dalam ketiga kategori, sehingga dapat dilakukan regresi logistik ordinal. Serta untuk data IPM yang dapat digunakan berasal dari tiga Provinsi yaitu Jawa Barat, Jawa Tengah dan Jawa Timur. Karena Provinsi Banten, DKI Jakarta dan D.I. Yogyakarta memiliki data sampel yang sedikit, maka tidak akan digunakan dalam penelitian ini. Untuk itu selanjutnya hanya akan digunakan data dari ketiga provinsi yaitu Jawa Barat, Jawa Tengah dan Jawa Timur.

##### a. Statistika Deskriptif

Pemodelan IPM di Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur tahun 2019 diawali dengan deskripsi dari variable penelitian. Deskripsi dari variable respon dari 100 Kabupaten/Kota di Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur yaitu Kabupaten/Kota yang memiliki IPM sedang sebanyak 40 kabupaten/kota, IPM tinggi sebanyak 50 kabupaten/kota dan IPM sangat tinggi sebanyak 10 kabupaten/kota. Selanjutnya untuk variable bebas disajikan pada Tabel 1.

**Tabel 1. Deskripsi Variabel bebas**

Variabel	Minimum	Maksimum	Rata-rata	Variansi
LPE ( $X_1$ )	0,14	7,14	5,334	1,174
TINGKAT KEMISKINAN ( $X_2$ )	2,07	20,71	9,562	14,322
SUMBER AIR LAYAK ( $X_3$ )	70,04	100	93,861	36,424
TPT ( $X_4$ )	0,95	10,28	5,084	4,840
TPAK ( $X_5$ )	61,31	79,55	68,111	12,373

b. Uji Multikolinearitas

Sebelum dilakukan analisis regresi logistik ordinal, maka perlu dilakukan pengujian multikolinearitas antar variable bebas untuk melihat kebebasan antar variable independent. Didapatkan nilai VIF keseluruhan untuk masing-masing variabel bebas kurang dari 2. Hal ini mengindikasikan bahwa tidak ada kasus multikolinieritas antara variabel bebas, sehingga semua variabel bebas bisa dimasukkan dalam model.

c. Model IPM berdasarkan Provinsi

i. Pengujian Serentak dan Parsial

Pengujian secara serentak digunakan untuk mengetahui kebermaknaan variabel secara bersama-sama. Hipotesis yang akan digunakan dalam uji ini adalah:  
Hipotesis:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_5 = 0$$

$$H_1: \text{Paling sedikit ada satu } \beta_k \neq 0, k=1,2,\dots,5$$

**Tabel 2 Nilai Statistik G dan p-value menurut Provinsi**

Uraian	Jabar	Jateng	Jatim
Statistik G	26,921	29,125	36,063
DF	5	5	5
p-value	0,000	0,000	0,000

Berdasarkan hasil perhitungan pada Tabel 2, dapat dilihat bahwa nilai p-value yang dihasilkan pada pengujian se-rentak di provinsi Jabar, Jateng dan Jatim  $< 0,1$  sehingga dapat dikatakan bahwa secara serentak model untuk ketiga provinsi layak / bisa di-terima. Selanjutnya, dilakukan pengujian parsial untuk mengetahui variable yang berpengaruh signifikan terhadap model pada ketiga provinsi Jawa Barat, Jawa tengah dan Jawa Timur.

Hipotesis:

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0, k=1,2,\dots,5$$

**Tabel 3. Uji Parsial (Wald)**

Penduga	Jawa barat		Jawa Tengah		Jawa Timur	
	Koef	pvalue	Koef	pvalue	Koef	pvalue
Konst 1	-1,405	0,945	1,289	0,961	18,544	0,338
Konst 2	5,556	0,788	6,416	0,809	24,171	0,215
LPE	0,487	0,476	1,208	0,075	-0,257	0,746
Miskin	-1,154	0,004	-0,453	0,018	-0,612	0,008
Air Bersih	0,174	0,080	0,253	0,157	0,164	0,217
TPT	-0,025	0,957	-0,660	0,062	1,075	0,064
TPAK	-0,152	0,541	-0,305	0,133	0,111	0,533

Berdasarkan Tabel 3, dapat diketahui bahwa variabel prediktor yang berpengaruh dengan taraf signifikan 0,1 terhadap nilai IPM di tiap provinsi berbeda.

Provinsi Jawa Barat dipengaruhi oleh variable Tingkat kemiskinan dan Sumber air bersih, Provinsi Jawa Tengah dipengaruhi oleh variable Laju Pertumbuhan ekonomi berdasarkan PDB harga konstan, Tingkat kemiskinan dan Tingkat pengangguran terbuka. Provinsi Jawa Timur dipengaruhi oleh variable Tingkat Kemiskinan dan Tingkat pengangguran terbuka.

ii. Uji Kesesuaian Model

Berdasarkan pengujian kesesuaian model didapatkan nilai p-value bernilai melebihi taraf signifikansi  $\alpha = 10\%$  yang berarti menunjukkan bahwa model di tiap-tiap provinsi sudah cukup memenuhi atau sesuai. Selanjutnya dilakukan pemodelan secara serentak dengan variabel yang telah berpengaruh secara signifikan baik secara serentak maupun parsial, Setelah didapatkan model terbaik, maka peluang tiap-tiap kategori di tiap-tiap provinsi juga bisa dihitung, sebagai berikut:

Provinsi Jawa Barat

$$\hat{\pi}_1(x) = \frac{\exp(7,243 + 1,105 X_2 - 0,181 X_3)}{1 + \exp(7,243 + 1,105 X_2 - 0,181 X_3)}$$

$$\hat{\pi}_2(x) = \frac{\exp(13,759 + 1,105 X_2 - 0,181 X_3)}{1 + \exp(13,759 + 1,105 X_2 - 0,181 X_3)} - \hat{\pi}_1(x)$$

$$\hat{\pi}_3(x) = \frac{1}{1 + \exp(13,759 + 1,105 X_2 - 0,181 X_3)}$$

Provinsi Jawa Tengah

$$\hat{\pi}_1(x) = \frac{\exp(1,131 - 1,395 X_1 + 0,472 X_2 + 0,246 X_4)}{1 + \exp(1,131 - 1,395 X_1 + 0,472 X_2 + 0,246 X_4)}$$

$$\hat{\pi}_2(x) = \frac{\exp(5,246 - 1,395 X_1 + 0,472 X_2 + 0,246 X_4)}{1 + \exp(5,246 - 1,395 X_1 + 0,472 X_2 + 0,246 X_4)} - \hat{\pi}_1(x)$$

$$\hat{\pi}_3(x) = \frac{1}{1 + \exp(5,246 - 1,395 X_1 + 0,472 X_2 + 0,246 X_4)}$$

Provinsi Jawa Timur

$$\hat{\pi}_1(x) = \frac{\exp(-3,640 + 0,635 X_2 - 1,050 X_4)}{1 + \exp(-3,640 + 0,635 X_2 - 1,050 X_4)}$$

$$\hat{\pi}_2(x) = \frac{\exp(1,848 + 0,635 X_2 - 1,050 X_4)}{1 + \exp(1,848 + 0,635 X_2 - 1,050 X_4)} - \hat{\pi}_1(x)$$

$$\hat{\pi}_3(x) = \frac{1}{1 + \exp(1,848 + 0,635 X_2 - 1,050 X_4)}$$

Berdasarkan hitungan peluang di atas dapat diperoleh hasil prediksi, sehingga kebenaran model logit ini dapat dilihat berdasar hasil pengklasifikasian antara prediksi dan observasi. Berdasarkan perhitungan ketepatan pengklasifikasian IPM pada dapat diketahui persentase ketepatan pengklasifikasian IPM terbesar berdasarkan model terbaik adalah model IPM untuk Provinsi Jawa Tengah yakni 82,85%, untuk model IPM Jawa Timur yakni 76,31%, untuk provinsi Jawa Barat sebesar 77,78%. Angka ini cukup baik karena ketepatan klasifikasi ketiga provinsi sudah lebih dari 50 persen, sehingga dapat disimpulkan bahwa model dianggap baik.

d. Pemodelan IPM Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur

Selanjutnya dilakukan pemodelan IPM dari ketiga provinsi di Pulau Jawa yaitu Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur untuk mengetahui faktor-faktor yang mempengaruhi IPM ketiga provinsi tersebut sekaligus. Setelah didapatkan estimasi parameter kemudian dilakukan pengujian-pengujian sebagai berikut:

i. Pengujian Serentak dan Parsial

Pengujian secara serentak digunakan untuk mengetahui kebermaknaan variabel secara bersamaan. Hipotesis yang akan digunakan dalam uji ini adalah :

Hipotesis:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_5 = 0$$

$$H_1: \text{Paling sedikit ada satu } \beta_k \neq 0, k=1,2,\dots,7$$

Berdasarkan hasil pengujian serentak didapatkan nilai  $G^2$  sebesar 76,047 dan p-value 0,000. Sehingga dapat dilihat bahwa nilai p-value yang dihasilkan pada pengujian se-mentak di ketiga provinsi kurang dari 0,1 sehingga dapat dikatakan bahwa secara serentak model layak / bisa di-terima. Selanjutnya, dilakukan pengujian parsial untuk mengetahui variabel yang berpengaruh signifikan terhadap model dengan dummy region pada 3 provinsi di Pulau Jawa.

Hipotesis:

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0, k=1,2,\dots,7$$

**Tabel 4. Uji Parsial (Wald)**

Penduga	3 Provinsi		
	Koef	Std. Error	pvalue
Konst 1	3,094	9,728	0,750
Konst 2	7,981	9,745	0,413
LPE	0,664	0,358	0,064
Miskin	-0,555	0,115	0,000
Air Bersih	0,143	0,058	0,013
TPT	-0,063	0,205	0,756
TPAK	-0,127	0,100	0,203
D1	1,479	1,033	0,152
D2	1,369	0,996	0,169

Berdasarkan Tabel 4, dapat diketahui bahwa variabel prediktor yang berpengaruh dengan taraf signifikan 0,1 terhadap nilai IPM di ketiga provinsi di Pulau Jawa dipengaruhi oleh variabel Laju pertumbuhan Ekonomi, variabel Tingkat kemiskinan, variabel Sumber air bersih.

ii. Uji Kesesuaian Model

Berdasarkan uji kesesuaian model didapatkan nilai  $\chi^2$  pada metode Deviance sebesar 117,975 nilai p-value bernilai melebihi  $\alpha$ . Hal ini menunjukkan bahwa model di ketiga provinsi dengan dummy variabel sudah cukup memenuhi atau sesuai.

Selanjutnya dilakukan pemodelan secara serentak dengan variabel yang telah berpengaruh secara signifikan baik secara serentak maupun parsial, dan didapatkan variabel yang berpengaruh signifikansi adalah variabel Laju pertumbuhan Ekonomi, variabel Tingkat kemiskinan, variabel Sumber air bersih.

Dikarenakan setelah dilakukan pemodelan IPM dengan variabel yang berpengaruh signifikan didapatkan hasil bahwa untuk variabel dummy 1 dan dummy tidak signifikan sehingga pemodelan untuk ketiga provinsi sekaligus tidak dapat dibedakan. Setelah didapatkan model terbaik, maka peluang tiap-tiap kategori di Ketiga Provinsi juga bisa dihitung, sebagai berikut:

$$\hat{\pi}_1(x) = \frac{\exp(16,114 + 0,657 X_1 - 0,468 X_2 + 0,188 X_3)}{1 + \exp(16,114 + 0,657 X_1 - 0,468 X_2 + 0,188 X_3)}$$

$$\hat{\pi}_2(x) = \frac{\exp(20,672 + 0,657 X_1 - 0,468 X_2 + 0,188 X_3)}{1 + \exp(20,672 + 0,657 X_1 - 0,468 X_2 + 0,188 X_3)} - \hat{\pi}_1(x)$$

$$\hat{\pi}_3(x) = 1 - \hat{\pi}_1(x) - \hat{\pi}_2(x) = \frac{1}{1 + \exp(20,672 + 0,657 X_1 - 0,468 X_2 + 0,188 X_3)}$$

Interpretasi model yang terbentuk dengan menggunakan nilai odds ratio untuk  $X_1$  laju pertumbuhan ekonomi sebesar 1,928 yang berarti bahwa laju pertumbuhan ekonomi PDB harga konstan memiliki peluang 1,928 lebih besar daripada variable lainnya dalam mempengaruhi tingkat IPM di ketiga provinsi. Nilai odds ratio untuk  $X_2$  tingkat kemiskinan sebesar 0,626 yang berarti bahwa tingkat kemiskinan memiliki peluang 0,626 lebih kecil daripada variable lainnya dalam mempengaruhi tingkat IPM di ketiga provinsi. Nilai odds ratio untuk  $X_3$  sumber air bersih layak sebesar 1,208 yang berarti bahwa sumber air bersih layak memiliki peluang 1,208 lebih besar daripada variable lainnya dalam mempengaruhi tingkat IPM di ketiga provinsi.

Hasil pengklasifikasian antara prediksi dan observasi. Berdasarkan perhitungan ketepatan pengklasifikasian IPM pada dapat diketahui persentase ketepatan pengklasifikasian IPM berdasarkan model terbaik untuk ketiga Provinsi yakni 73%. Angka ini cukup baik karena ketepatan klasifikasi sudah lebih dari 50 persen, sehingga dapat disimpulkan bahwa model dianggap baik.

## 5. KESIMPULAN

Kesimpulan yang didapat untuk menjawab permasalahan, berdasarkan analisis dan pembahasan dari penelitian ini untuk mengetahui pengaruh indikator yang diduga berpengaruh pada IPM di provinsi Jawa Timur, Jawa Tengah, Jawa Barat, yakni :

- a. Faktor yang berpengaruh pada IPM Jawa Barat adalah Tingkat kemiskinan dan Sumber air layak dengan ketepatan klasifikasi model 77,78%. Faktor yang berpengaruh pada IPM Jawa Tengah adalah Laju Pertumbuhan ekonomi berdasarkan PDB harga konstan, Tingkat kemiskinan dan Tingkat pengangguran terbuka dengan ketepatan klasifikasi model 82,85%. Faktor yang berpengaruh pada IPM Jawa Barat adalah Tingkat Kemiskinan dan Tingkat pengangguran terbuka dengan ketepatan klasifikasi model 76,31%.
- b. Faktor yang berpengaruh pada IPM dari ketiga Provinsi adalah Laju pertumbuhan Ekonomi, variable Tingkat kemiskinan, variabel Sumber air bersih dengan ketepatan klasifikasi model 73%.

## DAFTAR PUSTAKA

- Abreau, M. N. S *et al.* (2008), “ *Ordinal logistic regression models: application in quality of life studies*”, *Cadernos de Saude Publica*, Vol.24, suppl.
- Agresti, A. (2002), “*Categorical Data Analysis, 2nd edition*, John Willey and Sons, Inc.”, Hoboken, New Jersey.
- BPS (2008), *Indeks Pembangunan Manusia 2006-2007*. BPS, Jakarta.
- Diana, R. (2009). *Uji Kesamaan Vektor Parameter Model Linier Multivariat : Studi Kasus Faktor-faktor yang mempengaruhi IPM tahun 2007*. Jurusan Statistika-ITS, Surabaya.
- Fatimah, Citra. (2010). *Pemodelan IPM Provinsi Jawa Timur, Jawa Tengah, Jawa Barat Dan Sumatera Utara Dengan Metode Regresi Logistik Ordinal*. Jurusan Statistika-ITS, Surabaya.
- Hosmer, D. W. dan Lemeshow, S. (2000). *Applied Logistic Regression*. John Wiley and Son, New York.
- Lutfia, A. (2017). *Analisis Regresi Logistik Multinomial (studi kasus: Jenis pelanggaran Lalu Lintas Kota Surabaya)*. Jurusan Vokasi Statistika-ITS, Surabaya.
- Pentury, T, dkk (2016). *Analisis Regresi Logistik Ordinal (Studi kasus: Akreditasi SMA di Kota Ambon)*. Jurusan Matematika-Univeristas Pattimura, Ambon .