

**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)**

Ratih Nurmalasari¹, Dwi Ispriyanti², Sudarno³

¹Mahasiswa Departemen Statistika FSM Universitas Diponegoro

^{2,3}Staff Pengajar Departemen Statistika FSM Universitas Diponegoro

ABSTRACT

Human Development Index (HDI) is one of the most important indicator to observe another dimensions of human development. The HDI is a measurement for achievement levels of the quality of human development. This study analyze HDI in the Districts/Cities of Central Java in 2014. The Central Java's HDI data is categorized as low, medium, and high. The HDI presumed to be affected by many factors, such as high school participation rates, middle school graduates percentage, percentage of household with clean water access, numbers of health facility, open unemployment rate, and labour force participation rate. This study used the ordinal logistic regression and the ordinal probit regression as its statistical analysis method. The result showed that factors affecting HDI in the Districts/Cities of Central Java in 2014 are percentage of household with clean water access and numbers of health facility. To evaluate the performance of ordinal logistic regression and the ordinal probit regression, researcher uses classification accuracy and AIC. Based on research classification accuracy and AIC of each methods, the result showed that both the ordinal logistic regression and the ordinal probit regression has good result in analyzing factors affecting Human Development Index in the Districts/Cities of Central Java in 2014.

Keywords: HDI, Ordinal Logistic Regression, Ordinal Probit Regression, Classification Accuracy, AIC

1. PENDAHULUAN

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). Namun persoalannya adalah capaian pembangunan manusia secara parsial sangat bervariasi dimana beberapa aspek pembangunan tertentu berhasil dan beberapa aspek pembangunan lainnya gagal. Menurut BPS (2014), Indeks Pembangunan Manusia (IPM) merupakan salah satu indikator yang penting dalam melihat sisi lain dari pembangunan. IPM digunakan untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk).

Beberapa penelitian sebelumnya mengenai IPM, menyebutkan bahwa variabel yang berpengaruh terhadap IPM adalah Angka Partisipasi Sekolah (APS), jumlah sarana kesehatan, persentase rumah tangga dengan akses air bersih, tingkat partisipasi angkatan kerja (TPAK) dan PDRB per kapita. Data IPM yang digunakan pada penelitian ini 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. Selain menggunakan regresi logistik

ordinal, untuk data dengan variabel dependen berupa data kategori atau ordinal dan variabel independennya berupa data kuantitatif dapat menggunakan regresi probit ordinal.

Penelitian ini dilakukan untuk mengetahui faktor-faktor yang mempengaruhi IPM di Jawa tengah tahun 2014, serta mengetahui model terbaiknya menggunakan metode Regresi Logistik Ordinal dan Regresi Probit Ordinal.

2. TINJAUAN PUSTAKA

2.1 Regresi Logistik Ordinal

Menurut Hosmer dan Lemeshow (2000) regresi logistik ordinal merupakan salah satu metode statistika untuk menganalisis variabel respon (dependen) yang mempunyai skala data ordinal yang terdiri atas tiga kategori atau lebih. Variabel bebas (independen) yang dapat disertakan dalam model berupa data kategori atau kontinu yang terdiri atas dua variabel atau lebih.

Di dalam Agresti (2002) model untuk regresi logistik ordinal adalah model logit kumulatif (*cumulative logit models*). Pada model logit ini sifat ordinal dari respon Y dituangkan dalam peluang kumulatif. Misalkan variabel respon Y memiliki G buah kategori berskala ordinal dan \mathbf{x}_i menyatakan vektor yang terdiri dari p variabel pengamatan ke- i , $\mathbf{x}_i = [x_{i1} \ x_{i2} \ \dots \ x_{ip}]^T$ dengan $i = 1, 2, \dots, n$, maka model logit kumulatif dinyatakan sebagai berikut:

$$\text{logit}[P(Y_i \leq g | \mathbf{x}_i)] = \alpha_g + \mathbf{x}_i^T \boldsymbol{\beta}, \quad g = 1, 2, \dots, G-1 \quad (1)$$

dimana $P(Y_i \leq g | \mathbf{x}_i)$ adalah peluang kumulatif kurang dari atau sama dengan kategori ke- g jika diketahui \mathbf{x}_i , $\{\alpha_g\}$ merupakan parameter intersep dan memenuhi kondisi $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{G-1}$ dan $\boldsymbol{\beta} = [\beta_1 \ \beta_2 \ \dots \ \beta_p]^T$ merupakan vektor koefisien regresi yang bersesuaian dengan $x_1 \ x_2 \ \dots \ x_p$.

Di dalam Agresti (2002) logit kumulatif didefinisikan sebagai:

$$\text{logit}[P(Y_i \leq g | \mathbf{x}_i)] = \ln \left[\frac{P(Y_i \leq g | \mathbf{x}_i)}{1 - P(Y_i \leq g | \mathbf{x}_i)} \right], \quad g = 1, 2, \dots, G - 1$$

Maka model regresi logistik ordinal dapat dinyatakan sebagai berikut:

$$\text{logit}[P(Y_i \leq g | \mathbf{x}_i)] = \ln \left[\frac{P(Y_i \leq g | \mathbf{x}_i)}{1 - P(Y_i \leq g | \mathbf{x}_i)} \right] = \alpha_g + \mathbf{x}_i^T \boldsymbol{\beta}, \quad g = 1, 2, \dots, G - 1$$

Sehingga peluang untuk masing-masing kategori respon dapat dinyatakan sebagai berikut:

$$\pi_g(\mathbf{x}_i) = \frac{\exp(\alpha_g + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_g + \mathbf{x}_i^T \boldsymbol{\beta})} - \frac{\exp(\alpha_{g-1} + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_{g-1} + \mathbf{x}_i^T \boldsymbol{\beta})}, \quad g = 1, 2, \dots, G$$

2.2 Penaksiran Parameter Regresi Logistik Ordinal

Di dalam Agresti (2002) penaksiran parameter regresi logistik ordinal dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Jika diambil n sampel vektor variabel random Y_1, Y_2, \dots, Y_n , dengan $Y_i = [y_{i1} \ y_{i2} \ \dots \ y_{i,G}]^T$ berdistribusi multinomial dengan peluang hasil kategori ke- g adalah $\pi_g(\mathbf{x}_i)$, maka membentuk fungsi *likelihood* yaitu:

$$\ell(\boldsymbol{\theta}) = \prod_{i=1}^n \prod_{g=1}^G (\pi_g(\mathbf{x}_i))^{y_{ig}} \prod_{i=1}^n \prod_{g=1}^G \left[\frac{\exp(\alpha_g + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_g + \mathbf{x}_i^T \boldsymbol{\beta})} - \frac{\exp(\alpha_{g-1} + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_{g-1} + \mathbf{x}_i^T \boldsymbol{\beta})} \right]^{y_{ig}}$$

Menurut Hosmer dan Lemeshow (2000) prinsip dari metode MLE adalah mengestimasi vektor parameter $\boldsymbol{\theta} = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_{G-1} \ \beta_1 \ \beta_2 \ \dots \ \beta_p]^T$ dengan cara memaksimalkan fungsi *likelihood*. Untuk mempermudah perhitungan, maka dilakukan transformasi \ln pada fungsi *likelihood* sehingga terbentuk fungsi *ln-likelihood*, yaitu:

$$L(\theta) = \ln[\ell(\theta)] = \sum_{i=1}^n \sum_{g=1}^G y_{ig} \ln \left[\frac{\exp(\alpha_g + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_g + \mathbf{x}_i^T \boldsymbol{\beta})} - \frac{\exp(\alpha_{g-1} + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_{g-1} + \mathbf{x}_i^T \boldsymbol{\beta})} \right] \quad (2)$$

Jika variabel respon y_i pada persamaan (2) mempunyai 3 buah kategori ($G = 3$), maka fungsi *ln-likelihood* menjadi:

$$L(\theta) = \sum_{i=1}^n \left\{ y_{i1} \ln \left[\frac{\exp(\alpha_1 + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_1 + \mathbf{x}_i^T \boldsymbol{\beta})} \right] + y_{i2} \ln \left[\frac{\exp(\alpha_2 + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_2 + \mathbf{x}_i^T \boldsymbol{\beta})} - \frac{\exp(\alpha_1 + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_1 + \mathbf{x}_i^T \boldsymbol{\beta})} \right] \right. \\ \left. + y_{i3} \ln \left[1 - \frac{\exp(\alpha_2 + \mathbf{x}_i^T \boldsymbol{\beta})}{1 + \exp(\alpha_2 + \mathbf{x}_i^T \boldsymbol{\beta})} \right] \right\} \quad (3)$$

Menurut Hosmer dan Lemeshow (2000) estimasi parameter melalui metode MLE. Turunan parsial pertama dari fungsi *ln-likelihood* yang akan diestimasi merupakan fungsi yang nonlinier terhadap parameter. Estimasi parameter dari persamaan regresi yang nonlinier tidak mudah jika menggunakan MLE dan memerlukan metode yang bersifat iterasi, sehingga menurut Agresti (2002) metode iterasi yang digunakan adalah metode iterasi *Newton Raphson*.

2.3 Regresi Probit Ordinal

Regresi probit ordinal adalah model regresi yang dapat digunakan untuk menjelaskan hubungan antara variabel dependen yang merupakan variabel diskrit berskala ordinal dan variabel independen yang terdiri dari variabel kontinu, diskrit atau campuran antar keduanya.

Pemodelan regresi probit ordinal diawali dengan memperhatikan model berikut (Greene, 2000)

$$Y^* = \mathbf{x}^T \boldsymbol{\beta} + \varepsilon \quad (4)$$

dimana Y^* adalah variabel respon yang merupakan variabel diskrit, $\boldsymbol{\beta}$ adalah vektor parameter koefisien dengan $\boldsymbol{\beta} = [\beta_1 \ \beta_2 \ \dots \ \beta_p]^T$, \mathbf{x} adalah vektor variabel bebas, dengan $\mathbf{x} = [X_1 \ X_2 \ \dots \ X_p]^T$ dan ε adalah error yang diasumsikan berdistribusi $N(0, \sigma^2)$.

Berdasarkan persamaan (4) dilakukan transformasi ke dalam bentuk $Z = \frac{Y^* - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}$, dimana $Z \sim N(0,1)$ selanjutnya dilakukan pengkategorian terhadap Y^* secara ordinal yaitu untuk $Y^* \leq \gamma_1$ dikategorikan dengan $Y = 0$, untuk $\gamma_1 < Y^* \leq \gamma_2$ dikategorikan dengan $Y = 1$, ..., untuk $\gamma_{j-1} < Y^* \leq \gamma_j$ dikategorikan dengan $Y = j-1$, ..., untuk $Y^* > \gamma_j$ dikategorikan dengan $Y = j$, sehingga diperoleh model regresi probit ordinal sebagai berikut :

$$P(Y = 0) = P\left(Z \leq \frac{\gamma_1 - (\beta_0 + \mathbf{x}^T \boldsymbol{\beta})}{\sigma}\right) = P\left(Z \leq \frac{(\gamma_1 - \beta_0) - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}\right) = P\left(Z \leq \frac{\delta_1 - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}\right) = \Phi\left(\frac{\delta_1 - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}\right)$$

$$P(Y = 1) = P\left(\frac{(\gamma_1 - \beta_0) - \mathbf{x}^T \boldsymbol{\beta}}{\sigma} < Z \leq \frac{(\gamma_2 - \beta_0) - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}\right) = \Phi\left(\frac{\delta_2 - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}\right) - \Phi\left(\frac{\delta_1 - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}\right)$$

⋮

$$P(Y = j) = 1 - P(Y^* \leq \gamma_j) = 1 - P\left(Z \leq \frac{(\gamma_j - \beta_0) - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}\right) = 1 - \Phi\left(\frac{\delta_j - \mathbf{x}^T \boldsymbol{\beta}}{\sigma}\right)$$

dengan $Y = 0$ untuk kategori terendah dan $Y = j$ untuk kategori tertinggi dan $\Phi(\cdot)$ adalah fungsi distribusi kumulatif distribusi normal.

2.4 Estimasi Parameter Regresi Probit Ordinal

Estimasi parameter regresi probit ordinal menggunakan metode *Maximum Likelihood Estimation* (MLE). Berdasarkan model regresi probit yang diperoleh, jika

diambil sampel sebanyak n maka sampel randomnya adalah $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$ dimana $\mathbf{Y}_u = [Y_{0u} Y_{1u} Y_{2u} \dots Y_{j-1,u}]^T$ untuk $u = 1, 2, \dots, n$ dan \mathbf{Y}_u berdistribusi multinomial.

$\mathbf{Y}_u \sim M[1; P(Y = 0|\mathbf{x}_u), P(Y = 1|\mathbf{x}_u), P(Y = 2|\mathbf{x}_u), \dots, P(Y = j - 1|\mathbf{x}_u)]$

Sehingga dapat diperoleh fungsi ln likelihood sebagai berikut:

$$\ln L(\beta) = \sum_{u=1}^n y_{0u} \ln \phi\left(\frac{\delta_1 - \mathbf{x}^T \beta}{\sigma}\right) + \sum_{u=1}^n \left(1 - \sum_{i=0}^{j-1} y_{iu}\right) \ln \left(1 - \phi\left(\frac{\delta_j - \mathbf{x}^T \beta}{\sigma}\right)\right) + \sum_{u=1}^n \sum_{i=1}^{j-1} y_{iu} \left(\ln \left(\phi\left(\frac{\delta_{i+1} - \mathbf{x}^T \beta}{\sigma}\right) - \phi\left(\frac{\delta_i - \mathbf{x}^T \beta}{\sigma}\right)\right)\right)$$

Berdasarkan hasil penaksiran untuk parameter β dengan *Maximum Likelihood Estimation* tidak ditemukan bentuk yang *closed form*. Akibatnya penaksir parameter bagi β tidak bisa langsung diperoleh dan untuk mendapatkan penaksir *maximum likelihood* bagi β digunakan metode Newton-Raphson.

2.4 Efek Marginal

Menurut Greene (2000) untuk menginterpretasikan model regresi probit ordinal digunakan efek marginal (*marginal effect*) sebagai berikut:

$$\frac{\partial P(Y = 0)}{\partial x_k} = \phi\left(\frac{\delta_1 - \mathbf{x}^T \beta}{\sigma}\right) \left(\frac{-\beta_k}{\sigma}\right)$$

Persamaan tersebut menyatakan besarnya pengaruh variabel bebas X_k untuk $k=1,2, \dots, p$ terhadap $P(Y=0)$

$$\frac{\partial P(Y = 1)}{\partial x_k} = \phi\left(\frac{\delta_2 - \mathbf{x}^T \beta}{\sigma}\right) - \phi\left(\frac{\delta_1 - \mathbf{x}^T \beta}{\sigma}\right) \left(\frac{\beta_k}{\sigma}\right)$$

Persamaan tersebut menyatakan besarnya pengaruh variabel bebas X_k untuk $k=1,2, \dots, p$ terhadap $P(Y=1)$

$$\frac{\partial P(Y = j)}{\partial x_k} = \phi\left(\frac{\delta_j - \mathbf{x}^T \beta}{\sigma}\right) \left(\frac{\beta_k}{\sigma}\right)$$

Persamaan tersebut menyatakan besarnya pengaruh variabel bebas X_k untuk $k=1,2, \dots, p$ terhadap $P(Y=j)$

2.5 Uji Rasio Likelihood (Uji Parameter Keseluruhan)

Hipotesis yang digunakan adalah:

$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$ (secara bersama-sama variabel independen tidak mempengaruhi model)

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

Statistik Uji Rasio Likelihood : $G^2 = -2 \ln \left(\frac{L_1(\theta)}{L_2(\theta)}\right)$

dengan:

$L_1(\theta) = \text{fungsi likelihood tanpa variabel bebas}$

$L_2(\theta) = \text{fungsi likelihood dengan variabel bebas}$

Kriteria Uji

H_0 ditolak apabila nilai $G^2_{hit} > X^2_{(\alpha,p)}$ atau nilai signifikansi $< \alpha$

2.6 Uji Wald (Uji Parameter Secara Individu)

Hipotesis yang digunakan adalah:

$H_0 : \beta_k = 0$ (parameter tidak signifikan atau variabel bebas tidak memiliki hubungan yang kuat dengan variabel respon)

$H_1 : \beta_k \neq 0$, dengan $k = 1, 2, \dots, p$ (parameter signifikan atau variabel bebas memiliki hubungan yang kuat dengan variabel respon)

Statistik Uji:

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

Dengan $\hat{\beta}_k$ merupakan penaksir parameter β_k dan standar error $\hat{\beta}_k$ diperoleh dari:

$$SE(\hat{\beta}_k) = \sqrt{\widehat{var}(\hat{\beta}_k)}$$

Kriteria Uji:

H_0 ditolak apabila $W_k > X^2_{(\alpha,1)}$ atau nilai signifikansi $< \alpha$

2.7 Uji Kesesuaian Model (Goodness Of Fit)

Hipotesis yang digunakan adalah:

H_0 : Model sesuai (tidak ada perbedaan antara hasil observasi dengan hasil prediksi)

H_1 : Model tidak sesuai (ada perbedaan antara hasil observasi dengan hasil prediksi)

Statistik Uji :

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

dimana:

$\hat{\pi}_{ig} = \hat{\pi}_g(x_i)$: merupakan peluang observasi ke- i pada kategori ke- g

$df = J - (p + 1)$ dimana J merupakan jumlah kovariat, p = banyaknya parameter dalam model

Kriteria Uji :

H_0 ditolak jika $D > X^2_{(df;\alpha)}$ atau nilai signifikansi $< \alpha$

2.8 Ketepatan Klasifikasi

Ketepatan klasifikasi pada penelitian ini menggunakan APER (*Apparent Error Rate*). Nilai APER ini menunjukkan proporsi observasi yang salah diklasifikasikan oleh fungsi klasifikasi (Johnson dan Wichern, 2007). Nilai ketepatan klasifikasi dapat menggunakan $1 - \text{APER}$ (Johnson dan Wichern, 2007).

2.9 Akaike's Information Criterion (AIC)

Pemilihan model terbaik berdasarkan kriteria AIC dilakukan dengan memilih model yang memiliki nilai AIC terkecil. Rumus AIC yaitu sebagai berikut:

$$AIC(P^*) = \frac{-2 \ln L(P^*)}{n} + \frac{2p^*}{n}$$

Dengan $\ln L(P^*)$ adalah nilai maksimum likelihood yang mengandung p^* variabel prediktor, p^* adalah jumlah parameter β dimana $p^* = 0, 1, 2, \dots, p$ dan n adalah ukuran sampel.

2.10 Indeks Pembangunan Manusia

Indeks pembangunan manusia merupakan suatu ukuran yang digunakan untuk mengukur pencapaian pembangunan manusia di suatu wilayah. Meskipun tidak mengukur semua dimensi dari pembangunan manusia, namun IPM dinilai mampu mengukur dimensi pokok dari pembangunan manusia (BPS, 2013).

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

1. Angka Partisipasi Sekolah (APS) Tingkat SMA / MA

Angka Partisipasi Sekolah (APS) menunjukkan besaran penduduk usia sekolah yang sedang bersekolah (Melliana dan Zain, 2013).

2. Persentase penduduk yang tamat SMP atau sederajat

Pendidikan yang rendah membatasi kemampuan seseorang untuk mencari dan memanfaatkan peluang, dengan pendidikan yang tinggi tentunya akan mencetak Sumber Daya Manusia yang berkualitas baik dari segi spiritual, intelegensi dan kemampuan.

3. Persentase rumah tangga dengan akses 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. Banyaknya sarana kesehatan

Sarana kesehatan meliputi rumah sakit umum maupun swasta, puskesmas, dan puskesmas pembantu. Sarana kesehatan merupakan indikator untuk mengukur keberhasilan pembangunan dalam hal kesehatan (Melliani dan Zain, 2013).

6. Tingkat Partisipasi Angkatan Kerja (TPAK)

Menurut BPS Provinsi Jatim (2015) tingkat partisipasi angkatan kerja adalah persentase banyaknya angkatan kerja terhadap penduduk usia kerja.

3. METODOLOGI PENELITIAN

Sumber data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik Provinsi Jawa Tengah yaitu data hasil Survei Sosial Ekonomi Nasional (SUSENAS) Tahun 2014. 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 35 Kabupaten/Kota di Provinsi Jawa Tengah dan menggunakan 6 variabel bebas yang diduga mempengaruhi IPM. Data IPM dikategorikan menjadi 3 kategori, yaitu 1 = rendah ($IPM < 70$), 2 = sedang ($70 \leq IPM \leq 75$), dan 3 = tinggi ($IPM > 75$). Variabel prediktor yang akan digunakan pada penelitian ini adalah X_1 = Angka Partisipasi Sekolah (APS) SMA/MA, X_2 = persentase penduduk yang tamat SMP atau Sederajat, X_3 = persentase rumah tangga dengan akses air bersih, X_4 yaitu tingkat pengangguran terbuka (TPT), X_5 = banyaknya sarana kesehatan dan X_6 = tingkat partisipasi angkatan kerja (TPAK).

Tahapan-tahapan analisis yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Mengumpulkan data yang diperoleh dari Badan Pusat Statistik (BPS) yang digunakan dalam penelitian.
2. Mengkategorikan variabel dependen (1= $IPM < 70$), (2= $70 \leq IPM \leq 75$)), (3= $IPM > 75$)).
3. Melakukan deskriptif data sebagai gambaran pada data IPM Kabupaten/Kota di Jawa Tengah tahun 2014 serta faktor-faktor yang diduga memiliki pengaruh terhadap IPM.
4. Pembentukan model awal dari metode regresi logistik ordinal dan regresi probit ordinal pada data Indeks Pembangunan Manusia (IPM)
5. Melakukan uji rasio likelihood, uji wald, dan uji kesesuaian model terhadap model regresi logistik ordinal dan regresi probit ordinal.
6. Pembentukan model akhir dari regresi logistik ordinal dan regresi probit ordinal pada data Indeks Pembangunan Manusia (IPM)
7. Melakukan perhitungan kriteria model terbaik dengan ketepatan klasifikasi dan AIC menggunakan model akhir dari regresi logistik ordinal dan regresi probit ordinal.
8. Melakukan perbandingan hasil ketepatan klasifikasi dan AIC yang telah dihitung dengan metode regresi logistik ordinal dan regresi probit ordinal.

4. ANALISIS DAN PEMBAHASAN

4.1 Regresi Logistik Ordinal

A. Model Awal

$$\text{Logit } [P(Y_i \leq 1|X_i)] = 66,999 - 0,283 X_1 - 0,285 X_2 - 0,194X_3 - 0,184X_4$$

$$\text{Logit } [P(Y_i \leq 2|X_i)] = 73,132 - 0,283 X_1 - 0,285 X_2 - 0,194X_3 - 0,184X_4 + 0,145 X_5 - 0,430 X_6$$

B. Uji Rasio Likelihood

Tabel.1 Uji Rasio Likelihood

Uji Tahap ke-	G	Nilai Tabel Chi-Square	Keputusan
1	40,061	12,592	H ₀ ditolak
2	31,314	5,991	H ₀ ditolak

Berdasarkan pada Tabel 1 dapat dilihat untuk pengujian tahap pertama dan tahap kedua uji rasio *likelihood* (uji parameter secara keseluruhan) bahwa H₀ ditolak karena nilai $G > \chi^2_{(\alpha,p)}$, maka dapat disimpulkan bahwa paling sedikit ada satu variabel independen memberi pengaruh signifikan terhadap variabel dependen.

C. Uji Wald

Tabel 2. Uji Wald Tahap Pertama

Variabel	Wald	Sig.	Keputusan
X ₁	8,535	0,003	H ₀ ditolak
X ₂	0,409	0,523	H ₀ diterima
X ₃	2,359	0,125	H ₀ diterima
X ₄	0,114	0,736	H ₀ diterima
X ₅	6,511	0,011	H ₀ ditolak
X ₆	3,098	0,078	H ₀ diterima

Tabel 3. Uji Wald Tahap Kedua

Variabel	Wald	Sig.	Keputusan
X ₁	8,535	0,003	H ₀ ditolak
X ₅	6,511	0,011	H ₀ ditolak

Berdasarkan uji Wald tahap pertama dan tahap kedua yang telah dilakukan didapatkan hasil bahwa variabel yang signifikan terhadap model yaitu angka partisipasi sekolah (APS) tingkat SMA/MA (X₁) dan banyaknya sarana kesehatan (X₅).

D. Uji Kesesuaian Model

Berdasarkan uji kesesuaian model diperoleh nilai $Deviance = 32,548 < X^2_{(0,05;66)} = 85,9649$ atau nilai signifikansi = $1,000 > \alpha = 0,05$ maka H₀ diterima. Sehingga dapat disimpulkan bahwa model sesuai, sehingga tidak ada perbedaan antara hasil observasi dengan hasil prediksi.

E. Model Akhir

$$\text{Logit } [P(Y_i \leq 1|X_i)] = 8,756 - 0,208 X_1 + 0,112 X_5$$

$$\text{Logit } [P(Y_i \leq 2|X_i)] = 13,194 - 0,208 X_1 + 0,112 X_5$$

F. Ketepatan Klasifikasi

Tabel 4. APER Metode Regresi Logistik Ordinal

Observasi	Prediksi			Jumlah
	Kategori 1	Kategori 2	Kategori 3	
Kategori 1	18	3	0	21
Kategori 2	3	7	0	10
Kategori 3	0	1	3	4
Jumlah	21	11	3	35

Berdasarkan Tabel di atas dapat diperoleh nilai APER atau *error* yaitu $APER = \frac{3+0+3+0+0+1}{18+3+0+3+7+0+0+1+3} \times 100\% = 20\%$. Sehingga dapat disimpulkan bahwa nilai ketepatan klasifikasinya sebesar (1-APER) yaitu 80 %.

G. Akaike's Information Criterion (AIC)

Diperoleh nilai AIC untuk model regresi logistik ordinal. Dengan hasil yang diperoleh adalah:

$$AIC(P^*) = \frac{-2 \ln L(P^*)}{n} + \frac{2p^*}{n} = 1,158521$$

4.2 Regresi Probit Ordinal

A. Model Awal

$$Z_1 = 36,077 - 0,152X_1 - 0,139X_2 - 0,106X_3 - 0,069X_4 + 0,078X_5 - 0,235X_6$$

$$Z_2 = 39,243 - 0,152X_1 - 0,139X_2 - 0,106X_3 - 0,069X_4 + 0,078X_5 - 0,235X_6$$

B. Uji Rasio Likelihood

Tabel.5 Uji Rasio Likelihood

Uji Tahap ke-	G	Nilai Tabel Chi-Square	Keputusan
1	63,862	12,592	H ₀ ditolak
2	31,416	5,991	H ₀ ditolak

Berdasarkan pada Tabel 5 dapat dilihat untuk pengujian tahap pertama dan tahap kedua uji rasio *likelihood* (uji parameter secara keseluruhan) bahwa H₀ ditolak karena nilai $G > \chi^2_{(\alpha,p)}$, maka dapat disimpulkan bahwa paling sedikit ada satu variabel independen memberi pengaruh signifikan terhadap variabel dependen.

C. Uji Wald

Tabel 7. Uji Wald Tahap Pertama

Variabel	Wald	Sig.	Keputusan
X ₁	10,313	0,001	H ₀ ditolak
X ₂	0,337	0,561	H ₀ diterima
X ₃	2,478	0,115	H ₀ diterima
X ₄	0,055	0,814	H ₀ diterima
X ₅	7,830	0,005	H ₀ ditolak
X ₆	3,385	0,066	H ₀ diterima

Tabel 8. Uji Wald Tahap Kedua

Variabel	Wald	Sig.	Keputusan
X ₁	12,597	0,000	H ₀ ditolak
X ₅	9,386	0,002	H ₀ ditolak

Berdasarkan uji Wald tahap pertama dan tahap kedua yang telah dilakukan didapatkan hasil bahwa variabel yang signifikan terhadap model yaitu angka partisipasi sekolah (APS) tingkat SMA/MA (X₁) dan banyaknya sarana kesehatan (X₅).

D. Uji Kesesuaian Model

Berdasarkan uji kesesuaian model diperoleh nilai $Deviance = 32,446 < X^2_{(0,05;66)} = 85,9649$ atau nilai signifikansi = $1,000 > \alpha = 0,05$ maka H₀ diterima. Sehingga dapat disimpulkan bahwa model sesuai, sehingga tidak ada perbedaan antara hasil observasi dengan hasil prediksi.

E. Model Akhir

$$Z_1 = 4,890 - 0,116 X_1 + 0,061X_5$$

$$Z_2 = 7,356 - 0,116 X_1 + 0,061 X_5$$

F. Efek Marginal

Interpretasi model regresi probit ordinal dapat dilakukan berdasarkan efek marginal. Sebagai contoh berikut efek marginal Angka Partisipasi Sekolah (APS) tingkat SMA/MA terhadap Indeks Pembangunan Manusia (IPM):

$$\frac{\partial P(Y = 1)}{\partial x_1} = -\beta_1(\phi(Z_1)) = 0,116(\phi(4,890 - 0,116 X_1 + 0,061X_5)) = 0,092934$$

Artinya perubahan nilai angka partisipasi sekolah SMA/MA sebesar satu unit akan berpengaruh terhadap probabilitas kabupaten Cilacap mendapatkan IPM dengan kategori rendah sebesar 0,092934.

$$\begin{aligned} \frac{\partial P(Y = 2)}{\partial x_1} &= \beta_1(\phi(Z_2) - \phi(Z_1)) \\ &= -0,116(\phi(7,356 - 0,116X_1 + 0,061X_5) - \phi(4,890 - 0,116X_1 + 0,061X_5)) \\ &= -0,02301 \end{aligned}$$

Artinya perubahan nilai angka partisipasi sekolah SMA/MA sebesar satu unit akan berpengaruh terhadap probabilitas kabupaten Cilacap mendapatkan IPM kategori sedang sebesar -0,02301.

$$\frac{\partial P(Y = 3)}{\partial x_1} = \beta_1(\phi(Z_2)) = -0,116(\phi(7,356 - 0,116 X_1 + 0,061 X_5)) = -0,11595$$

Artinya perubahan nilai angka partisipasi sekolah SMA/MA sebesar satu unit akan berpengaruh terhadap probabilitas kabupaten Cilacap mendapatkan IPM dengan kategori tinggi sebesar -0,11595.

G. Ketepatan Klasifikasi

Tabel 6. APER Metode Regresi Logistik Ordinal

Observasi	Prediksi			Jumlah
	Kategori 1	Kategori 2	Kategori 3	
Kategori 1	18	3	0	21
Kategori 2	3	7	0	10
Kategori 3	0	1	3	4
Jumlah	21	11	3	35

Berdasarkan Tabel di atas dapat diperoleh nilai APER atau *error APER* = $\frac{3+0+3+0+0+1}{18+3+0+3+7+0+0+1+3} \times 100\% = 20\%$. Sehingga dapat disimpulkan bahwa nilai ketepatan klasifikasinya sebesar (1-APER) yaitu 80 %.

H. Akaike's Information Criterion (AIC)

Diperoleh nilai AIC untuk model regresi logistik ordinal. Dengan hasil yang diperoleh adalah:

$$AIC(P^*) = \frac{-2 \ln L(P^*)}{n} + \frac{2p^*}{n} = 1,155612$$

4.3 Perbandingan Metode Regresi Logistik Ordinal dengan Regresi Probit Ordinal

Tabel 7. Perbandingan Metode Regresi Logistik Ordinal dan Regresi Probit Ordinal

Model	Ketepatan Klasifikasi	AIC
Regresi Logistik Ordinal	80 %	1,158521
Regresi Probit Ordinal	80 %	1,155612

Berdasarkan Tabel 7 diperoleh bahwa nilai ketepatan klasifikasi yang dihasilkan oleh regresi logistik ordinal memiliki hasil yang sama besar dengan regresi probit ordinal, sedangkan nilai AIC dari regresi logistik ordinal dan regresi probit ordinal memiliki hasil relatif sama. Sehingga dapat disimpulkan bahwa metode regresi logistik ordinal dan metode regresi probit ordinal sama baiknya untuk menganalisis faktor-faktor yang mempengaruhi Indeks Pembangunan Manusia (IPM) di Jawa Tengah Tahun 2014.

5. KESIMPULAN

Berdasarkan hasil dan pembahasan mengenai Indeks Pembangunan Manusia (IPM) di Jawa Tengah tahun 2014 menggunakan regresi logistik ordinal dan regresi probit ordinal, diperoleh kesimpulan sebagai berikut:

1. Indeks Pembangunan Manusia (IPM) di Jawa Tengah dipengaruhi oleh faktor Angka Partisipasi Sekolah (APS) SMA/MA dan banyaknya sarana kesehatan.
2. Nilai ketepatan klasifikasi metode regresi logistik ordinal adalah sebesar 80 % dan nilai ketepatan klasifikasi metode regresi probit ordinal adalah sebesar 80 %. Nilai AIC metode regresi logistik ordinal sebesar 1,158521 dan nilai AIC regresi probit ordinal sebesar 1,155612.
3. Dari hasil perbandingan ketepatan klasifikasi dan AIC antara metode regresi logistik ordinal dengan regresi probit ordinal dapat disimpulkan bahwa metode regresi

logistik ordinal sama baiknya dengan metode regresi probit ordinal dalam menganalisis faktor-faktor yang mempengaruhi Indeks Pembangunan Manusia (IPM) di Jawa Tengah tahun 2014.

DAFTAR PUSTAKA

- Agresti, A. 2002. *Categorical Data Analysis*, Second Edition. John Willey and Sons New York.
- BPS. 2014. *Indeks Pembangunan Manusia 2013*. Jakarta: BPS.
- BPS. 2015. *Indeks Pembangunan Manusia 2014*. Jakarta: BPS.
- Dewi, S.R. 2014. Pemodelan Metode Geographically Weighted Ordinal Logistic Regression dengan Fungsi Pembobot Adaptive Gaussian Kernel, Adaptive Bisquare Kernel, dan Adaptive Tricube Kernel, *Jurnal Mahasiswa Statistik*, Vol. 2, No.2, Hal. 429-432.
- Greene, W.H. 2000. *Econometrics Analysis*, Sixth Edition. Prentice Hall: New Jersey.
- Gujarati, D.N. 2012. *Dasar-dasar Ekonometrika*, Buku 2 Edisi 4. Jakarta: Salemba Empat.
- Hosmer, D.W., dan Lemeshow. 2000. *Applied Logistic Regression*. USA : John Wiley and Sons.
- Johnson, R. A. dan Wichern, D. W. 2007. *Applied Multivariate Statistical Analysis*, Sixth Edition. New Jersey: Prentice Hall.
- Melliana, A. dan Zain, I. 2013. Analisis Faktor yang Mempengaruhi Indeks Pembangunan Manusia di Kabupaten/Kota Provinsi Jawa Timur dengan Menggunakan Regresi Panel, *Jurnal Sains dan Seni POMITS*, Vol. 2, No. 2, Hal. 237-242.
- Pradita, R.N. 2015. Pemodelan Faktor-Faktor yang mempengaruhi Indeks Pembangunan Manusia Kabupaten/Kota di Jawa Timur Menggunakan *Geographically Weighted Ordinal Logistic Regression*, *Jurnal Gaussian*, Vol. 4, No. 3, Hal. 639 – 650.
- Rachmasita, K. 2011. Pemodelan Regresi Probit Ordinal Pada Persentase Sekolah Terklasifikasi Hitam Menurut Pola Jawaban UN, *Jurnal Sains dan Seni POMITS*, Vol. 4, No. 1, Hal. 103 - 120.