JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI e-ISSN: 2614-8811
Published by Departement of Mathematics, Hasanuddin University, Indonesia
https://journal.unhas.ac.id/index.php/jmsk/index
Vol. 20, No. 3, May 2024, pp. 538-551
DOI: 10.20956/j.v20i3.32356

Classification Of Country Status In 2022 Based On Social Indicators With Ordinal Logistic Regression

Klasifikasi Status Negara Tahun 2022 Berdasarkan Indikator Sosial Menggunakan Regresi Logistik Ordinal

Sugha Faiz Al Maula[1]*, Alfi Nur Nitasari[2]*, Mochamad Rasyid Aditya Putra[3]*, Maelcardino Christopher Justin[4]*, Salma Bethari Andjani Sumarto[5]*, Suliyanto[6]*, Toha Saifudin[7]*

*** *Department of Mathematics, Faculty of Science and Technology, University of Airlangga, Indonesia*
***Email**:[1] sugha.faiz.al-2021@fst.unair.ac.id , [2]alfi.nur.nitasari-2021@fst.unair.ac.id, [3]mochamad.rasyid.aditya-2021@fst.unair.ac.id, [4]maelcardino.christopher.justin-2021@fst.unair.ac.id, [5]salma.bethari.andjani-2021@fst.unair.ac.id, [6]suliyanto@fst.unair.ac.id, [7]tohasaifudin@fst.unair.ac.id*

Received: 18 December 2023, revised: 26 February 2024, accepted: 27 February 2024

Abstract

This research examines the classification of country status in 2022 by applying ordinal logistic regression on various social indicators including education, health and economic. The urgency of the research is to know the country determine factors with specific factors in the form of research variables that can be useful for policy makers, unlike the existing classification which is only divided based on GDP per capita or HDI score only. By dividing 3 country status classes, namely not developed, developing and developed countries using the world bank classification baseline, the accuracy results were obtained at 72,5% but there were several variables that were not significant. After re-modelling, the accuracy was found increased to 76.4% with the odds ratio results for the minimum wage variable being 42,32 in the high class compared to the middle class and 11,66 for the middle class compared to the lower class, which means that the higher the minimum wage tends to be classify countries as developed countries. Another variable that has significance level is the birth rate with an odds ratio of 0,71 in the high and middle classes and 0.89 in the middle and lower classes comparison, which shows that this variable has a negative effect because the odds ratio is <1, which means that the higher the birth rate tends to make the country will be classified as a non-developed country.

Keywords: *Country Classification, Ordinal Logistic Regression, Social Indicator*

Abstrak

Penelitian ini mengkaji klasifikasi status negara pada tahun 2022 dengan menerapkan regresi logistik ordinal pada berbagai indikator sosial termasuk pendidikan, kesehatan, dan pendapatan.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Urgensi dari penelitian yaitu mengetahui faktor klasifikasi negara dengan faktor spesifik berupa variabel penelitian yang dapat berguna bagi pemangku kebijakan, tidak seperti klasifikasi yang ada dengan hanya membagi berdasarkan PDB perkapita atau skor IPM saja. Dengan membagi 3 kelas status negara yaitu tidak maju, berkembang, dan negara maju menggunakan *baseline* klasifikasi world bank, didapatkan hasil akurasi sebesar 72,5% namun memiliki beberapa variabel yang tidak signifikan. Setelah dilakukan pemodelan ulang didapatkan akurasi yang meningkat menjadi 76,4% dengan hasil rasio *odds* pada variabel upah minimum sebesar 42,32 pada kelas tinggi dibandingkan dengan kelas menengah dan 11,66 untuk kelas menengah dibandingkan dengan kelas bawah yang berarti semakin tinggi upah minimum cenderung mengklasifikasikan negara menjadi negara maju. Adapun variabel lain yang berpengaruh adalah angka kelahiran dengan rasio *odds* sebesar 0,71 pada kelas tinggi dan menengah dan 0,89 pada kelas menengah dan kelas bawah yang menunjukkan bahwasanya variabel tersebut berpengaruh negatif karena rasio *odds* <1 yang berarti semakin tinggi angka kelahiran cenderung membuat negara akan terklasifikasikan sebagai negara tidak maju.

Kata kunci: *Klasifikasi Negara, Regresi Logistik Ordinal, Indikator Sosial*

1. PENDAHULUAN

Perkembangan negara di dunia saat ini tentu tidak selalu berjalan beriringan, beberapa negara sudah maju karena sumber daya alamnya yang melimpah atau karena riset pengembangan teknologi di negara tersebut, sedangkan yang lainnya masih menghadapi konflik di dalam negeri maupun isu sosial lainnya. Alasan tersebut menjadi salah satu sebab klasifikasi negara menjadi penting, seperti klasifikasi milik International Monetary Fund (IMF) atau World Bank yang meminjamkan pinjaman lunak kepada negara yang berhak menerima serta negara maju sebagai pendonornya Perkembangan negara di dunia saat ini tentu tidak selalu berjalan beriringan, beberapa negara sudah maju karena sumber daya alamnya yang melimpah atau karena riset pengembangan teknologi di negara tersebut, sedangkan yang lainnya masih menghadapi konflik di dalam negeri maupun isu sosial lainnya. Alasan tersebut menjadi salah satu sebab klasifikasi negara menjadi penting, seperti klasifikasi milik International Monetary Fund (IMF) atau World Bank yang meminjamkan pinjaman lunak kepada negara yang berhak menerima serta negara maju sebagai pendonornya [7].

Selain karena alasan bantuan berupa pinjaman lunak, pengklasifikasian negara juga bertujuan untuk menyederhanakan kompleksivitas negara di dunia yang sangat beragam menjadi kelompok negara yang relatif homogen. Hal tersebut dapat menjadi proses pembangunan yang dapat dicontoh dan dibandingkan dengan negara lain yang sebanding [12].

Klasifikasi yang digunakan adalah tiga tingkatan yaitu negara maju, berkembang, serta tidak berkembang. Adapun karakteristik dari negara tidak berkembang diantaranya adalah negara tersebut cenderung hanya fokus kepada satu sumber daya alam atau tidak didiversifikasi dengan baik. Selain itu negara tidak berkembang cenderung pula fokus pada pertanian saja dengan produktivitas tenaga kerja yang rendah. Indikator selanjutnya adalah rendahnya tingkat pendidikan serta infrastruktur fisik yang belum memadai [1].

Adapun usaha dari negara untuk menjadi negara maju, dapat dilakukan dengan baik oleh Korea Selatan. Negara tersebut setelah mengalami perang saudara termasuk negara termiskin di dunia. Namun, setelah fase tersebut mereka bertranformasi hingga kini menjadi salah satu kekuatan ekonomi di Asia. Hal yang dilakukan antara lain menguatkan kebijakan ekonomi yang tepat dan koheren, inovasi pada teknologi dan sistemnya, serta mendorong Perusahaan besar dalam negeri untuk bersaing di pasar global [9].

Sementara itu kasus negara lain yang belum berhasil menaikkan tingkat kesejahteraan negaranya juga berusaha meraih dengan cara baru seperti Indonesia memiliki kebijakan diantaranya hilirisasi produk dan menjualnya dengan nilai tambah yang berarti tidak mengekspor dalam bentuk bahan mentah. Sayangnya, usaha tersebut digugat oleh negara-negara maju yang

manufakturnya sudah maju seperti Uni Eropa menggugat kebijakan tentang larangan ekspor bahan mentah. Uni Eropa bereaksi dengan mengajukan keluhan di WTO (World Trade Organization), yang menyatakan bahwa ekspor ban Indonesia telah membatasi akses produsen terhadap nikel secara tidak adil [13]. Permasalahan tersebut dapat menjadi contoh aplikatif dari manfaat mengklasifikasikan status negara mengingat Indonesia merupakan negara berkembang, maka usaha untuk menjadi negara maju perlu dilakukan dan salah satunya adalah dengan hilirisasi yang berkonsekuensi pada larangan ekspor bahan mentah.

Penelitian kali ini mengangkat tentang indikator sosial yang akan diklasifikasi menggunakan analisis regresi logistik ordinal karena nilai respon lebih dari dua yaitu negara maju, berkembang, dan tidak berkembang berdasarkan penilaian dari IMF dan PBB. Selain mendapatkan persamaan regresi logistik untuk setiap kelasnya, klasifikasi juga dapat memetakan negara tersebut sesuai kategori dari data yang diberikan serta melihat skor akurasi dari proses klasifikasi. Penelitian ini juga dapat menjadi acuan kedepannya dan mengambil langkah yang tepat guna bertransformasi menjadi negara maju pada variabel spesifik yang berpengaruh, karena klasifikasi yang tersedia hanya mengklasifikasikan berdasarkan PDB perkapita atau IPM dari hasil riset tanpa mengetahui secara jelas pengaruh aspek lain seperti kependudukan atau aspek ketenagakerjaan yang dapat diintervensi oleh pemerintah dan pada akhirnya dapat meningkatkan nilai PDB perkapita atau IPM.

2. TINJAUAN PUSTAKA

2.1 Model Logistik Ordinal

Regresi logistik ordinal merupakan salah satu metode statistika untuk menganalisis variabel respon yang memiliki skala ordinal yang terdiri dari tiga kategori atau lebih. Variabel prediktor pada model logistik ordinal dapat bersifat kategorik atau kontinu [11]. Pada data berpasangan $(Y, \boldsymbol{X})$, dengan Y merupakan variabel respon berskala ordinal dengan q kategori dan $\boldsymbol{X} = (X_1, X_2, \dots, X_p)$ merupakan suatu vektor dari variabel prediktor atau *covariates*. Misalkan $\pi_j(\boldsymbol{X}) = P(Y = j|\boldsymbol{X})$ merupakan peluang nilai $Y = j$ jika $\boldsymbol{X}$ diketahui untuk $j = 1,2, \dots, q$ yang dituliskan sebagai berikut.

$$\pi_j = \gamma_j - \gamma_{j-1} \tag{2.1}$$

dan $\gamma_j(\boldsymbol{X}) = P(\boldsymbol{Y} \geq j|\boldsymbol{X})$ adalah peluang kumulatif dari variabel respon kategori j untuk $j = 1,2, \dots, q$. Variabel respon Y berskala ordinal berhubungan dengan vektor peluang $\boldsymbol{\pi} = \left(\pi_1(\boldsymbol{X}), \pi_2(\boldsymbol{X}), \dots, \pi_q(\boldsymbol{X})\right)$ dan dinyatakan dalam model logistik ordinal sebagai berikut.

$$g\left(\gamma_j(\boldsymbol{X})\right) = \theta_j + \sum_{k=1}^{p} \beta_k X_k\,, \qquad j = 1,2, \dots, q-1 \tag{2.2}$$

Dengan

$$g\left(\gamma_j(\boldsymbol{X})\right) = \ln\left(\frac{\gamma_j(\boldsymbol{X})}{1-\gamma_j(\boldsymbol{X})}\right) \quad \text{adalah fungsi link logit} \tag{2.3}$$

2.2 Estimasi Model Logistik Ordinal

Dari $g\left(\gamma_j(\boldsymbol{X})\right) = \theta_j + \sum_{k=1}^{p} \beta_k X_k$, diperoleh peluang kumulatif sebagai berikut.

$$\gamma_j(\boldsymbol{X}) = P(Y \geq j|\boldsymbol{X}) = \frac{\exp\left(\theta_j + \sum_{k=1}^{p} \beta_k X_k\right)}{1 + \exp\left(\theta_j + \sum_{k=1}^{p} \beta_k X_k\right)} \tag{2.4}$$

Dari Persamaan (2.4) diperoleh Persamaan (2.5) sebagai berikut.

$$\gamma_j(\boldsymbol{X}) = \sum_{s=1}^{j} \pi_s(\boldsymbol{X}), \quad j = 1,2, \dots, q \quad \text{dan } \gamma_q(\boldsymbol{X}) = 1 \tag{2.5}$$

Misalkan n_i merupakan banyak ulangan observasi ke-i dan n_{ij} merupakan banyak observasi nilai $Y_i = j$ yang muncul pada nilai *covariate* $\boldsymbol{X_i}$, untuk $j = 1,2, \dots, q$ maka $[Y_i, i = 1,2, \dots, n]$

merupakan variabel random independen yang berdistribusi multinomial yang dinotasikan $Y_i \sim multinomial(n_i, \pi_i)$ dengan $n_i = \sum_{j=1}^{q} n_{ji}$, $\sum_{j=1}^{q} \pi_{ji} = 1$ dan $\boldsymbol{\pi_i} = (\pi_{1i}, \pi_{2i}, \dots, \pi_{qi})$.

Untuk mengestimasi parameter dengan model regresi logistik ordinal diperlukan peluang kumulatif, sehingga fungsi *likelihood* dapat dinyatakan sebagai perkalian $q-1$ kategori. Fungsi peluang bersama dari $(Y_1, Y_2, \dots, Y_n)$ adalah sama dengan perkalian n fungsi multinomial, dengan fungsi kepadatan peluang sebagai berikut [2].

$$p(Y = y) = p(Y_1 = y_1, \dots, Y_k = Y_k) = \frac{n!}{y_1! \, y_2! \dots y_k!} p_1^{y_1} \dots p_k^{y_k} \tag{2.6}$$

Metode pendugaan parameter yang dapat digunakan pada regresi logistik ordinal diantaranya adalah dengan metode kemungkinan maksimum yang dapat dilakukan jika diasumsikan saling bebas antara amatan yang satu dengan yang lain. Metode ini dapat dilakukan jika. Fungsi kemungkinan-nya dapat dinyatakan sebagai [5]:

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\beta}) &= \prod_{i=1}^{n} \pi_1(\boldsymbol{X_i})^{n_{1i}} \pi_2(\boldsymbol{X_i})^{n_{2i}} \pi_3(\boldsymbol{X_i})^{n_{3i}} \dots \pi_q(\boldsymbol{X_i})^{n_{qi}} \\ &= \prod_{i=1}^{n} \gamma_1(\boldsymbol{X_i})^{n_{1i}} [\gamma_2(\boldsymbol{X_i}) - \gamma_1(\boldsymbol{X_i})]^{n_{2i}} \dots \left[1 - \gamma_{q-1}(\boldsymbol{X_i})\right]^{n_{qi}} \end{aligned} \tag{2.7}$$

Berdasarkan Persamaan (2.7) diperoleh fungsi log *likelihood* $\ell = \ln L(\theta, \beta)$ sebagai berikut.

$$\begin{aligned} \ell &= \sum_{i=1}^{n} \ln\left\{\gamma_1(\boldsymbol{X_i})^{n_{1i}} [\gamma_2(\boldsymbol{X_i}) - \gamma_1(\boldsymbol{X_i})]^{n_{2i}} \dots \left[1 - \gamma_{q-1}(\boldsymbol{X_i})\right]^{n_{qi}}\right\} \\ &= \prod_{i=1}^{n} \left\{n_{1i} \ln \gamma_1(\boldsymbol{X_i}) + n_{2i} \ln[\gamma_2(\boldsymbol{X_i}) - \gamma_1(\boldsymbol{X_i})] + \dots + n_{qi} \ln\left[1 - \gamma_{q-1}(\boldsymbol{X_i})\right]\right\} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{q} n_{ji} \ln\left[\gamma_j(\boldsymbol{X_i}) - \gamma_{j-1}(\boldsymbol{X_i})\right] \\ &= \sum_{i=1}^{n} \sum_{j=1}^{q} n_{ji} \ln\left[\frac{e^{\theta_j + \boldsymbol{X_i\beta}}}{1 + e^{\theta_j + \boldsymbol{X_i\beta}}} - \frac{e^{\theta_{j-1} + \boldsymbol{X_i\beta}}}{1 + e^{\theta_{j-1} + \boldsymbol{X_i\beta}}}\right] \end{aligned} \tag{2.8}$$

Dari fungsi *log-likehood* tersebut, diperoleh turunan $\ln L(\boldsymbol{\theta}, \boldsymbol{\beta})$ terhadap $\boldsymbol{\theta}$ dan $\boldsymbol{\beta}$ sebagai berikut.

$$\begin{aligned} \frac{\partial \ell}{\partial \theta_1} &= \sum_{i=1}^{n} \left\{\left(\frac{n_{1i}}{1 + e^{\theta_1 + \boldsymbol{X_i\beta}}}\right) + n_{2i} \left(-\frac{e^{\theta_1}}{e^{\theta_2} - e^{\theta_1}} - \frac{e^{\theta_1 + \boldsymbol{X_i\beta}}}{1 + e^{\theta_1 + \boldsymbol{X_i\beta}}}\right)\right\} \\ \frac{\partial \ell}{\partial \theta_2} &= \sum_{i=1}^{n} \left\{n_{2i} \left(\frac{e^{\theta_2}}{e^{\theta_2} - e^{\theta_1}} - \frac{e^{\theta_2 + \boldsymbol{X_i\beta}}}{1 + e^{\theta_2 + \boldsymbol{X_i\beta}}}\right) + n_{2i} \left(\frac{e^{\theta_2}}{e^{\theta_3} - e^{\theta_2}} - \frac{e^{\theta_2 + \boldsymbol{X_i\beta}}}{1 + e^{\theta_2 + \boldsymbol{X_i\beta}}}\right)\right\} \\ \vdots \quad \dots \qquad & \qquad \vdots \end{aligned} \tag{2.9}$$

Turunan kedua dari fungsi *log-likelihood* berikut didapatkan turunan $\ln L(\boldsymbol{\theta}, \boldsymbol{\beta})$ terhadap $\boldsymbol{\theta}$ dan $\boldsymbol{\beta}$ sebagi berikut.

$$\frac{\partial^2 \ell}{\partial \theta_1^2} = \sum_{i=1}^{n} \left\{\frac{n_{1i} e^{\theta_1 + \boldsymbol{X_i\beta}}}{\left(1 + e^{\theta_1 + \boldsymbol{X_i\beta}}\right)^2} + n_{2i} \left(-\frac{e^{\theta_1} e^{\theta_2}}{\left(e^{\theta_2} - e^{\theta_1}\right)^2} - \frac{e^{\theta_1 + \boldsymbol{X_i\beta}}}{\left(1 + e^{\theta_1 + \boldsymbol{X_i\beta}}\right)^2}\right)\right\} \tag{2.10}$$

$\vdots$

Secara umum untuk $u = 2, \dots, q-1$ diperoleh

$$\begin{aligned} \frac{\partial^2 \ell}{\partial \theta_u^2} &= \sum_{i=1}^{n} \left\{-\left(n_{ui} + n_{(u+1)i}\right) \frac{e^{\theta_u + \boldsymbol{X_i\beta}}}{\left(1 + e^{\theta_u + \boldsymbol{X_i\beta}}\right)^2} - n_{ui} \frac{e^{\theta_u} e^{\theta_{u-1}}}{e^{\theta_u} - e^{\theta_{u-1}}} - \right. \\ &\quad \left. \left(n_{(u+1)i}\right) \frac{e^{\theta_u} e^{\theta_{u-1}}}{\left(e^{\theta_{u+1}} - e^{\theta_u}\right)^2}\right\} \\ &= \sum_{i=1}^{n} \left\{n_{21} \frac{e^{\theta_1} e^{\theta_2}}{\left(e^{\theta_2} - e^{\theta_1}\right)^2}\right\} \end{aligned} \tag{2.11}$$

$$\begin{array}{lcl} \vdots & \cdots & \vdots \\ \dfrac{\partial^2 \ell}{\partial\theta_u \partial\theta_{u-1}} & = & \sum_{i=1}^{n}\left\{n_{ui}\dfrac{e^{\theta_u}e^{\theta_{u-1}}}{\left(e^{\theta_u}-e^{\theta_{u-1}}\right)^2}\right\} \end{array}$$

$$\begin{array}{lcl} \dfrac{\partial^2 \ell}{\partial\theta_1 \partial\theta_j} & = & 0, \text{untuk } j = 3, \ldots, q-1 \\ \vdots & & \end{array} \tag{2.12}$$

Secara umum diperoleh

$$\begin{array}{lcl} \dfrac{\partial^2 \ell}{\partial\theta_u \partial\theta_j} & = & 0, |u-j| \geq 2 \\ \dfrac{\partial^2 \ell}{\partial\theta_1 \partial\beta_j} & = & \sum_{i=1}^{n}\left\{-\dfrac{n_{1i}X_{ji}e^{\theta_1+\boldsymbol{X_i\beta}}}{\left(1+e^{\theta_1+\boldsymbol{X_i\beta}}\right)^2}+n_{21}\dfrac{X_{ji}e^{\theta_1+\boldsymbol{X_i\beta}}}{\left(1+e^{\theta_1+\boldsymbol{X_i\beta}}\right)^2}\right\}, j = 1, \ldots, k \\ \dfrac{\partial^2 \ell}{\partial\theta_u \partial\beta_j} & = & \sum_{i=1}^{n}\left\{-\left(n_{ui}+n_{(u+1)i}\right)\dfrac{X_{ji}e^{\theta_u+\boldsymbol{X_i\beta}}}{\left(1+e^{\theta_u+\boldsymbol{X_i\beta}}\right)^2}\right\}, u = 2, \ldots, q-1 \\ \dfrac{\partial \ell}{\partial\beta_j} & = & \begin{array}{l}\sum_{i=1}^{n}\left\{n_{1i}\left(X_{ji}-\dfrac{X_{ji}e^{\theta_1+\boldsymbol{X_i\beta}}}{1+e^{\theta_1+\boldsymbol{X_i\beta}}}\right)+n_{21}\left(X_{ji}-\dfrac{X_{ji}e^{\theta_1+\boldsymbol{X_i\beta}}}{1+e^{\theta_1+\boldsymbol{X_i\beta}}}-\dfrac{X_{ji}e^{\theta_2+\boldsymbol{X_i\beta}}}{1+e^{\theta_2+\boldsymbol{X_i\beta}}}\right)+\right. \\ \left.\cdots+n_{qi}\left(-\dfrac{X_{ji}e^{\theta_{q-1}+\boldsymbol{X_i\beta}}}{1+e^{\theta_{q-1}+\boldsymbol{X_i\beta}}}\right)\right\}, j = 1,2,\ldots,k \end{array} \\ \dfrac{\partial^2 \ell}{\partial\beta_u \partial\beta_j} & = & \begin{array}{l}\sum_{i=1}^{n}\left\{n_{1i}\left(-\dfrac{X_{ui}X_{ji}e^{\theta_1+\boldsymbol{X_i\beta}}}{1+e^{\theta_1+\boldsymbol{X_i\beta}}}\right)+n_{21}\left(X_{ji}-\dfrac{X_{ui}X_{ji}e^{\theta_1+\boldsymbol{X_i\beta}}}{\left(1+e^{\theta_1+\boldsymbol{X_i\beta}}\right)^2}-\right.\right. \\ \left.\dfrac{X_{ui}X_{ji}e^{\theta_2+\boldsymbol{X_i\beta}}}{\left(1+e^{\theta_2+\boldsymbol{X_i\beta}}\right)^2}\right)+\cdots+n_{qi}\left(-\dfrac{X_{ui}X_{ji}e^{\theta_{q-1}+\boldsymbol{X_i\beta}}}{\left(1+e^{\theta_{q-1}+\boldsymbol{X_i\beta}}\right)^2}\right)\Bigg\}, j \neq 1,2,\ldots,k \end{array} \end{array} \tag{2.13}$$

Untuk mengestimasi vektor parameter $\boldsymbol{\theta}' = (\theta_1, \theta_2, \ldots, \theta_{q-1})$ dan $\boldsymbol{\beta}' = (\beta_1, \beta_2, \ldots, \beta_p)$ digunakan metode *Newton-Raphsone* sebagai berikut.

$\boldsymbol{\varphi}^{(i+1)} = \boldsymbol{\varphi}^{(i)} - \left(\boldsymbol{H}^{(i)}\right)^{-1}\boldsymbol{r}^{(i)}$ dengan $\boldsymbol{r}^{(i)} = \frac{\partial \ell}{\partial \boldsymbol{\varphi}^{(i)}}, i = 0,1, \ldots$ dan matriks Hessian $\boldsymbol{H} = \left(\frac{\partial^2 \ell}{\partial \boldsymbol{\varphi}\partial \boldsymbol{\varphi}^{\boldsymbol{T}}}\right)$ dengan $\boldsymbol{\varphi} = (\boldsymbol{\theta}', \boldsymbol{\beta}')'$. Estimasi vektor parameter diperoleh pada saat konvergen, yaitu ketika nilai maks $\left|\boldsymbol{\varphi}^{(i+1)} - \boldsymbol{\varphi}^{(i)}\right| < \delta$, dengan δ merupakan bilangan positif yang kecil [11].

2.3 Ukuran Kebaikan Model

Untuk menentukan kebaikan model regresi logistik ordinal dapat diukur menggunakan berbagai cara seperti *confusion matrix*, skor akurasi, dan rasio *odds*.

Confusion matrix merupakan salah satu teknik klasifikasi mengenai aktual dan prediksi yang dilakukan dengan sistem klasifikasi. *Confusion matrix* dapat digunakan untuk mengevaluasi kinerja model regresi logistik ordinal. Baris mewakili kategori sebenarnya, dan kolom mewakili kategori yang diprediksi. Elemen diagonal mewakili jumlah prediksi yang benar, dan elemen di luar diagonal mewakili jumlah prediksi yang salah [6]. *Confusion matrix* dapat digunakan untuk menghitung berbagai metrik kinerja seperti akurasi, presisi, perolehan, dan skor F1 untuk menentukan klasifikasi, akurasi, sensitivitas, dan spesifikasi. Metrik ini dapat dihitung untuk setiap kategori secara terpisah atau dapat dirata-ratakan di seluruh kategori.

Skor akurasi pengklasifikasian dikatakan baik ketika skor akurasi semakin mendekati dengan 100%. Dalam model logit ordinal, nilai *Odd* pada kategori $Y \leq j$ didefinisikan sebagai berikut [11].

$$Odd = \frac{P(Y \leq j|\boldsymbol{X})}{P(Y > j|\boldsymbol{X})} = \exp\left(\theta_j + \boldsymbol{X\beta}\right); j = 1,2, \ldots, q-1 \tag{2.14}$$

Berdasarkan nilai *Odd* berikut, didefinisikan nilai Rasio *Odds* (OR) dari variabel prediktor X_k pada kategori $Y \leq j; j = 1,2, \dots, q-1$ sebagai berikut.

$$OR_k = \frac{\frac{P(Y \leq j | X_k = x_{k-1})}{P(Y > j | X_k = x_{k-1})}}{\frac{P(Y \leq j | X_k = x_k)}{P(Y > j | X_k = x_k)}} = \exp(\beta_k); k = 1,2, \dots p \quad (2.15)$$

Nilai OR menggambarkan untuk setiap kenaikan variabel prediktor X_k sebesar satu unit mengakibatkan kenaikan atau penurunan *Odd* pada kategori $Y \leq j$ dengan $j = 1,2, \dots q-1$ secara multiplikatif sebesar $\exp(\beta_k)$.

3. METODE PENELITIAN

3.1 Data dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data Global Country Information Dataset 2023 yang merupakan data sekunder yang diperoleh dari *website* kaggle.com dan divalidasi di laman resmi *world bank* yang berisi indikator yang dimiliki negara-negara dunia dalam berbagai aspek. Variabel yang digunakan dalam penelitian disajikan pada Tabel 3.1 sebasgai berikut.

Tabel 3.1. Variabel Penelitian

Kode	Nama Variabel	Satuan	Tipe Data
Y	Status Negara 1: Negara Tidak Berkembang 2: Negara Berkembang 3: Negara Maju	-	Diskrit
X_1	Angka Kelahiran	Jumlah	Kontinu
X_2	Angka Kesuburan	Jumlah	Kontinu
X_3	Kematian Bayi	Jumlah	Kontinu
X_4	Upah Minimum	Jumlah	Kontinu
X_5	Partisipasi Angkatan Kerja	Persen	Kontinu
X_6	Pendapatan Pajak	Persen	Kontinu
X_7	Tingkat Pengangguran	Persen	Kontinu

3.2 Analisis Data

Metode yang digunakan dalam penelitian ini merupakan salah satu metode regresi parametrik dengan regresi logistik ordinal. Dalam penelitian ini, analisis dilakukan dengan menggunakan software python. Adapun langkah-langkah untuk melakukan analisis dalam penelitian ini adalah sebagai berikut.

1. Melakukan analisis deskriptif data
2. Modeling (mencari estimator β_j) menyusun dalam persamaan regresi logistik ordinal
3. Splitting data

   Data keseluruhan yang digunakan berjumlah 167 dengan pembagian data sebagai berikut.

Tabel 3.2. Splitting Data

Kelas	Kategori Negara	*Training*	*Testing*	Total
1	Tidak Berkembang	51	21	72
2	Berkembang	30	13	43
3	Maju	35	17	52

Total	116	51	167

Berdasarkan Tabel 3.2, diketahui negara dengan kategori 1 atau negara tidak berkembang memiliki jumlah data yang lebih banyak daripada negara dengan kategori 2 dan 3 yaitu negara berkembang dan negara maju.

4. Pembentukan model regresi logistik ordinal dari seluruh variabel prediktor yang berpengaruh dengan variabel respon sesuai Persamaan (3.1) berikut.

$$\gamma_j(\boldsymbol{X}) = P(Y \geq j|\boldsymbol{X}) = \frac{\exp\left(\theta_j + \sum_{k=1}^{p} \beta_k X_k\right)}{1 + \exp\left(\theta_j + \sum_{k=1}^{p} \beta_k X_k\right)} \quad (3.1)$$

5. Uji serentak

   Pengujian ini dilakukan untuk mengetahui ada tidaknya pengaruh variabel prediktor terhadap variable respon secara bersama-sama. Dalam pengujian serentak, uji signifikansi model dapat digunakan *likelihood ratio test*.

6. Uji parsial

   Pengujian ini dilakukan untuk mengetahui ada tidaknya pengaruh variable prediktor terhadap variable respon secara parsial atau sendiri-sendiri. Pengujian ini dilakukan dengan menggunakan statistic uji normal baku.

7. Performansi atau ukuran kebaikan model

   Ukuran kebaikan model dapat diukur dengan skor akurasi menggunakan *confusion matrix* yang nantinya dapat dihitung akurasi, presisi, perolehan, dan skor F1 untuk menentukan klasifikasi. Ukuran kebaikan model juga dapat diukur menggunakan Pseudo R-squared untuk mengukur seberapa baik model dapat menjelaskan variasi dalam data.

8. Melakukan interpretasi terhadap model regresi logistik ordinal yang diperoleh

   Interpretasi model regresi logistic ordinal dilakukan dengan melihat parameter yang dihasilkan model. Parameter tersebut dapat diinterpretasikan dengan menggunakan rasio *odds*. Jika nilai rasio *odds* > 1, maka peluang terjadinya perubahan variabel respon akan semakin besar dengan adanya peningkatan pada variabel prediktor. Sebaliknya jika nilai rasio *odds* < 1, maka peluang terjadinya perubahan variabel respon akan semakin kecil dengan adanya peningkatan pada variabel prediktor.

4. HASIL DAN PEMBAHASAN

4.1 Pemodelan dengan Regresi Logistik Ordinal

4.1.1 Mendapatkan Estimator $\beta_{(j)}$

Model regresi logistik ordinal memiliki estimator $\beta_{(j)}$ untuk tiap koefisien tiap variabel dan hanya tersedia sebanyak 2 kelas, karena kelas ke-3 (negara maju) merupakan *baseline* model pada persamaan. Data diacak dan dibagi menjadi training dan testing dengan rasio 70:30. Dengan fungsi penghubung logit serta bantuan *software* python dan Persamaan (2.7) hingga (2.13) maka didapatkan koefisien untuk kelas pertama dan kedua sebagai berikut:

Tabel 4.1. Koefisien Kelas

Variabel	Koefisien	
	2	3
Angka Kelahiran	-0,54	-1,03
Angka Kesuburan	2,37	3,14
Kematian Bayi	-0,13	-0,26
Upah Minimum	0,71	2,33
Partisipasi Angkatan Kerja	0,09	0,18

Pendapatan Pajak	-0,09	-0,23
Tingkat Pengangguran	0,45	0,44

4.1.2 Persamaan Model

Berdasarkan Tabel 4.1 dan Persamaan (3.1), diperoleh model persamaan untuk tiap kelas adalah sebagai berikut dengan probabilitas terbesar untuk setiap subjek adalah kategori kelasnya.

1. Persamaan regresi untuk kelas 1 = negara tidak berkembang

$$\begin{aligned} \pi_3(\boldsymbol{X}) &= P(\boldsymbol{Y} \geq \boldsymbol{3}|\boldsymbol{X}) \\ &= \frac{\exp(\theta_3+\sum_{k=1}^{p}\beta_k X_k)}{1+\exp(\theta_3+\sum_{k=1}^{p}\beta_k X_k)} \\ &= \frac{\exp(-0{,}54X_1+2{,}37X_2-0{,}13X_3+0{,}71X_4+0{,}09X_5-0{,}09X_6+0{,}45X_7)}{1+\exp(-0{,}54X_1+2{,}37X_2-0{,}13X_3+0{,}71X_4+0{,}09X_5-0{,}09X_6+0{,}45X_7)} \end{aligned} \quad (4.1)$$

2. Persamaan regresi untuk kelas 2 = negara berkembang

$$\begin{aligned} \pi_2(\boldsymbol{X}) &= \gamma_2(\boldsymbol{X}) - \pi_3(\boldsymbol{X}) \\ &= \frac{\exp(\theta_2+\sum_{k=1}^{p}\beta_k X_k)}{1+\exp(\theta_2+\sum_{k=1}^{p}\beta_k X_k)} - \frac{\exp(\theta_3+\sum_{k=1}^{p}\beta_k X_k)}{1+\exp(\theta_3+\sum_{k=1}^{p}\beta_k X_k)} \\ &= \frac{\exp(-1{,}03X_1+3{,}14X_2-0{,}26X_3+2{,}33X_4+0{,}18X_5-0{,}23X_6+0{,}44X_7)}{1+\exp(-1{,}03X_1+3{,}14X_2-0{,}26X_3+2{,}33X_4+0{,}18X_5-0{,}23X_6+0{,}44X_7)} - \\ &\quad \frac{\exp(-0{,}54X_1+2{,}37X_2-0{,}13X_3+0{,}71X_4+0{,}09X_5-0{,}09X_6+0{,}45X_7)}{1+\exp(-0{,}54X_1+2{,}37X_2-0{,}13X_3+0{,}71X_4+0{,}09X_5-0{,}09X_6+0{,}45X_7)} \end{aligned} \quad (4.2)$$

3. Persamaan regresi untuk kelas 3 = negara maju

$$\begin{aligned} \pi_1(\boldsymbol{X}) &= 1 - P(\boldsymbol{Y} \geq 2|\boldsymbol{X}) \\ &= 1 - \frac{\exp(\theta_2+\sum_{k=1}^{p}\beta_k X_k)}{1+\exp(\theta_2+\sum_{k=1}^{p}\beta_k X_k)} \\ &= 1 - \frac{\exp(-1{,}03X_1+3{,}14X_2-0{,}26X_3+2{,}33X_4+0{,}18X_5-0{,}23X_6+0{,}44X_7)}{1+\exp(-1{,}03X_1+3{,}14X_2-0{,}26X_3+2{,}33X_4+0{,}18X_5-0{,}23X_6+0{,}44X_7)} \end{aligned} \quad (4.3)$$

4.2 Hasil Evaluasi Model

4.2.1 Uji Parsial dan Serentak

Untuk menguji parameter model secara parsial dapat digunakan statistik uji normal baku dengan hipotesis sebagai berikut.

$H_0 : \beta_j = 0$

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

Secara asimtotik estimator maksimum *likelihood* $\hat{\boldsymbol{\varphi}} \sim \boldsymbol{N}(\boldsymbol{\varphi}, [I(\boldsymbol{\varphi})]^{-1})$, sehingga diperoleh $var(\hat{\beta}_j)$ adalah elemen diagonal ke j dari matriks $[I(\boldsymbol{\varphi})]^{-1}$. Akibatnya diperoleh simpangan baku dari estimator $\hat{\beta}_j$ adalah $s(\hat{\beta}_j) = \sqrt{var(\hat{\beta}_j)}$. Untuk menguji hipotesis tersebut digunakan statistik uji Wald dengan

$$Z_j = \frac{\hat{\beta}_j}{s(\hat{\beta}_j)}; j = 1,2,\dots,p \quad (4.4)$$

Statistik uji Z_j secara asimtotik berdistribusi normal baku. Daerah kritis untuk menguji hipotesis tersebut dengan tingkat signifikansi α adalah H_0 ditolak jika $|Z_j| > Z_{\frac{\alpha}{2}}$. Untuk sampel berukuran besar diperoleh interval kepercayaan sebesar $(1-\alpha)100\%$ untuk β_j adalah

$$\hat{\beta}_j \pm z_{\frac{\alpha}{2}} S(\hat{\beta}_j) \quad (4.5)$$

Adapun dalam pengujian serentak, uji signifikansi model dapat digunakan *likelihood ratio test*, dengan hipotesis yang diuji sebagai berikut.

$H_0 : \boldsymbol{\beta} = 0$

$H_1 : \boldsymbol{\beta} \neq 0$

Dengan daerah kritis tolak H_0 jika $P - Value < \alpha$ (0,05).

Berdasarkan Persamaan (4.4) dan (4.5) serta analisis dengan *Python*, diperoleh output untuk uji parsial dan uji serentak sebagai berikut.

Tabel 4.2. Output Uji Parsial dan Uji Serentak

Kategori Negara=2	Coef	$P > \|z\|$
Angka Kelahiran	−0,54	0,086
Angka Kesuburan	2,37	0,253
Kematian Bayi	−0,13	0,036
Upah Minimum	0,71	0,375
Partisipasi Angkatan Kerja	0,09	0,025
Pendapatan Pajak	−0,09	0,131
Tingkat Pengangguran	0,45	0,003
Kategori Negara=3	Coef	$P > \|z\|$
Angka Kelahiran	−1,03	0,061
Angka Kesuburan	3,14	0,526
Kematian Bayi	−0,26	0,205
Upah Minimum	2,33	0,025
Partisipasi Angkatan Kerja	0,18	0,015
Pendapatan Pajak	−0,23	0,040
Tingkat Pengangguran	0,44	0,041

4.2.2 Confusion Matrix

Confusion matrix atau biasa dikenal dengan *error matrix* biasa digunakan untuk klasifikasi di dalam statistik. Matriks tersebut berisi tabel yang menggambarkan performa dari model dalam mengklasifikasikan data *testing* dengan kelas yang ada, tiap baris pada matriks mewakili nilai kelas prediksi dari model sedangkan kolom menggambarkan kelas yang sesungguhnya dari data [3]. Adapun hasil *confusion matrix* dari model adalah sebagai berikut:

Tabel 4.3. *Confusion Matrix Logistic Regression*

	Predicted			Total
Actual	1	2	3	
1	15	6	1	22
2	5	6	1	12
3	0	1	16	17
Total	20	13	18	51

4.2.3 Perhitungan Metrik Klasifikasi

Ukuran kebaikan model digunakan untuk mengevaluasi sejauh mana model statistik cocok dengan data yang diobservasi. Karena model digunakan untuk klasifikasi, ukuran kebaikan model dilihat dari nilai akurasi untuk mengukur sejauh mana model benar-benar mengklasifikasikan data dengan benar. Nilai akurasi didapatkan dari:

$$accuracy = \frac{TP_1 + TP_2 + TP_3}{N} = \frac{15 + 6 + 16}{51} = 0{,}72549$$

Pada analisis ini, diperoleh nilai akurasi senilai 72,55% yang mana nilai tersebut cukup tinggi yang menandakan bahwa klasifikasi data dianggap baik. Selanjutnya metrik yang lain yaitu

precision dan *recall*. *Precision* adalah rasio prediksi positif yang benar dengan hasil keseluruhan dari prediksi positif, dengan kata lain, *precision* mengukur seberapa tepat model dalam memprediksi kelas positif. Adapun *Recall* adalah rasio prediksi positif yang benar dengan hasil keseluruhan dari kelas positif, dengan kata lain, *recall* mengukur seberapa lengkap model dalam memprediksi kelas positif. Selain itu perhitungan tersebut terdapat nilai yang menyeimbangkan antara kemampuan *recall* dan *precision* dalam model yang disebut dengan *F1-Score.*

Tabel 4.4. Rangkuman *Metrik Logistic Regression*

	Precission	*Recall*	*F1-Score*	*Support*
1	0,75	0,68	0,71	22
2	0,46	0,50	0,48	12
3	0,89	0,94	0,91	17
Accuracy			0,73	51

Pada perhitungan tersebut, diperoleh skor untuk tiap-tiap kelas pada 3 metrik yaitu *recall, precision, dan F1-Score.* Didapatkan bahwasanya model sudah cukup baik dalam memprediksi kelas 1 dan 3 namun masih kurang dalam memprediksi kategori kelas 2.

4.2.4 Rasio *Odds*

Dalam model logit ordinal didefinisikan *Odd* pada kategori $Y \leq j$ seperti pada Persamaan (2.14) sehingga diperoleh nilai rasio *odds* (OR) dari variabel prediktor X_k pada kategori $Y \leq j; j = 1,2, \dots, q-1$ sesuai dengan Persamaan (2.15) yang menggambarkan untuk setiap kenaikan variabel prediktor X_k sebesar satu unit mengakibatkan kenaikan atau penurunan *Odd* pada kategori $Y \leq j$ dengan $j = 1,2, \dots q-1$ secara multiplikatif sebesar $\exp(\beta_k)$. Berdasarkan Persamaan (2.15) serta analisis dengan *Python*, diperoleh rasio *odds* sebagai berikut.

Tabel 4.5. Rasio *Odds*

Variabel	Rasio *Odds* Kelas	
	$\frac{P(Y=2)}{P(Y=1)}$	$\frac{P(Y=3)}{P(Y=2)}$
Angka Kelahiran	0,58	0,36
Angka Kesuburan	10,70	23,03
Kematian Bayi	0,88	0,77
Upah Minimum	2,04	10,30
Partisipasi Angkatan Kerja	1,09	1,20
Pendapatan Pajak	0,91	0,79
Tingkat Pengangguran	1,56	1,55

4.2.5 *Re-modelling*

Dalam model dengan 7 variabel di atas, terdapat beberapa variabel yang tidak signifikan pada kelas 3 dan kelas 2. Maka dari itu, penulis melakukan pemodelan ulang dan didapatkan hasil sebagai berikut:

Tabel 4.6. Output Uji Parsial dan Uji Serentak

Kategori Negara=2	Coef	$P > \|z\|$
Angka Kelahiran	−0,11	0,000
Upah Minimum	2,45	0,000
Kategori Negara=3	Coef	$P > \|z\|$

Angka Kelahiran	−0,33	0,000
Upah Minimum	3,74	0,000

Berdasarkan Tabel 4.6 dan Persamaan (3.1), diperoleh re-modelling persamaan untuk tiap kelas adalah sebagai berikut dengan probabilitas terbesar untuk setiap subjek adalah kategori kelasnya.

1. Persamaan regresi untuk kelas 1 = negara tidak berkembang

$$\begin{aligned} \pi_3(\boldsymbol{X}) &= P(\boldsymbol{Y} \geq \mathbf{3}|\boldsymbol{X}) \\ &= \frac{\exp\left(\theta_3+\sum_{k=1}^{p}\beta_k X_k\right)}{1+\exp\left(\theta_3+\sum_{k=1}^{p}\beta_k X_k\right)} \\ &= \frac{\exp(-0{,}11X_1+2{,}45X_4)}{1+\exp(-0{,}11X_1+2{,}45X_4)} \end{aligned} \tag{4.6}$$

2. Persamaan regresi untuk kelas 2 = negara berkembang

$$\begin{aligned} \pi_2(\boldsymbol{X}) &= \gamma_2(\boldsymbol{X}) - \pi_3(\boldsymbol{X}) \\ &= \frac{\exp\left(\theta_2+\sum_{k=1}^{p}\beta_k X_k\right)}{1+\exp\left(\theta_2+\sum_{k=1}^{p}\beta_k X_k\right)} - \frac{\exp\left(\theta_3+\sum_{k=1}^{p}\beta_k X_k\right)}{1+\exp\left(\theta_3+\sum_{k=1}^{p}\beta_k X_k\right)} \\ &= \frac{\exp(-0{,}33X_1+3{,}74X_4)}{1+\exp(-0{,}33X_1+3{,}74X_4)} - \frac{\exp(-0{,}11X_1+2{,}45X_4)}{1+\exp(-0{,}11X_1+2{,}45X_4)} \end{aligned} \tag{4.7}$$

3. Persamaan regresi untuk kelas 3 = negara maju

$$\begin{aligned} \pi_1(\boldsymbol{X}) &= 1 - P(\boldsymbol{Y} \geq 2|\boldsymbol{X}) \\ &= 1 - \frac{\exp\left(\theta_2+\sum_{k=1}^{p}\beta_k X_k\right)}{1+\exp\left(\theta_2+\sum_{k=1}^{p}\beta_k X_k\right)} \\ &= 1 - \frac{\exp(-0{,}33X_1+3{,}74X_4)}{1+\exp(-0{,}33X_1+3{,}74X_4)} \end{aligned} \tag{4.8}$$

Tabel 4.7. *Confusion Matrix Logistic Regression*

	Predicted			Total
Actual	1	2	3	
1	16	5	0	21
2	6	7	0	13
3	0	1	16	17
Total	22	13	16	51

Berdasarkan Tabel 4.7 maka didapatkan skor akurasi sebesar 76,47% yang berarti mengalami penaikan setelah membuang variabel yang tidak signifikan

$$accuracy = \frac{TP_1 + TP_2 + TP_3}{N} = \frac{16 + 7 + 16}{51} = 0{,}7647$$

Tabel 4.8. Rangkuman *Metrik Logistic Regression Re-modelling*

	Precission	*Recall*	*F1-Score*	*Support*
1	0,73	0,76	0,74	21
2	0,54	0,54	0,54	13
3	1,00	0,94	0,97	17
Accuracy			0,76	51

Berdasarkan Tabel 4.8 maka didapatkan skor *precision* dan *recall* yang baik pada kelas tinggi dan rendah, namun masih kecil pada kelas menengah. Hal tersebut disebabkan negara

berkembang memiliki banyak karakteristik dan masalah yang berbeda-beda pada beberapa faktor sedangkan negara tidak maju memiliki semua faktor yang berjalan buruk dan negara maju semua faktor berjalan baik.

Tabel 4.9. Rasio *Odds*

Variabel	Rasio *Odds* Kelas $\frac{P(Y=2)}{P(Y=1)}$	$\frac{P(Y=3)}{P(Y=2)}$
Angka Kelahiran	0,89	0,71
Upah Minimum	11,66	42,32

Berdasarkan Tabel 4.9 didapatkan rasio *odds* pada variabel yang signifikan saja dan menggambarkan hasil pengklasifikasian dengan baik. Dengan aturan apabila rasio *odds* >1 maka variabel tersebut akan cenderung membuat negara diklasifikan sebagai negara maju dan sebaliknya.

4.3 Interpretasi dari Klasifikasi

Model dapat diinterpretasi dengan baik setelah melakukan perhitungan rasio *odds*. Perbandingan model untuk tiap kelas bersifat multipikatif dengan *baseline model* adalah kelas sebelumnya. Variabel angka kelahiran pada rasio *odds* kelas 2 menunjukkan jika 1 poin angka kelahiran bertambah maka potensi kecenderungan menurun adalah 0,89 dari kelompok sedang menuju ke rendah. Artinya angka kelahiran yang tinggi cenderung dimiliki oleh kelompok rendah. Adapun rasio *odds* kelas 3 menunjukkan jika 1 poin angka kelahiran bertambah maka potensi kecenderungan menurun adalah 0,71 dari kelompok tinggi ke kelompok sedang karena nilai rasio *odds* < 1. Hal tersebut sesuai dengan penelitian bahwasanya tingginya dan sulit dikontrolnya angka kelahiran akan menurunkan pertumbuhan ekonomi suatu negara [8]. Model yang didapatkan sesuai dengan kasus angka kelahiran dimana tingginya angka kelahiran dominan terjadi pada negara berpendapatan menengah kebawah dan berpendapatan rendah [10] pada model didapatkan rasio *odds* < 1 yang mengindikasikan angka kelahiran kelas-2 (negara menengah) lebih kecil dari kelompok negara tidak maju dan negara maju lebih kecil daripada negara menengah. Rentang yang baik dari negara maju untuk angka kelahiran adalah pada rentang 5,7-15,7 kelahiran per 1000 penduduk [4], sebagai perbandingan Indonesia berada diangka 16,4 pada 2023.

Adapun variabel kedua yang berpengaruh secara signifikan adalah variabel upah minimum dengan rasio *odds* 11,66 pada negara berkembang dibandingkan dengan tidak maju dan 42,32 pada negara maju dibandingkan dengan negara berkembang yang berarti bagi upah minimum semakin tinggi maka akan cenderung membawa negara tersebut diklasifikasikan sebagai negara maju. Secara PDB perkapita negara maju memiliki PDB perkapita minimal 13200$ secara rata-rata. Maka dari itu, guna menuju negara berpenghasilan tinggi, kebijakan upah minimum harus secara bertahap meningkat ke arah tersebut. Salah satu usaha untuk menuju kearah tersebut adalah dengan cara mengurangi pekerja sektor informal. Hal tersebut terkait dengan semakin besarnya prevalensi lapangan kerja informal diasosiasikan dengan rendahnya keterbelakangan perekonomian, semakin besarnya tingkat korupsi di sektor publik, semakin kecilnya pemerintahan dan semakin rendahnya tingkat intervensi negara untuk melindungi pekerja dari kemiskinan [14].

5. SIMPULAN

Berdasarkan analisis yang telah dilakukan, dapat diperoleh kesimpulan bahwa:

1. Berdasarkan pemodelan menggunakan regresi logistik ordinal didapatkan akurasi model sebear 72,55% dan meningkat menjadi 76,47% setelah mengeluarkan variabel yang tidak signifikan. Hal tersebut berarti secara umum model dapat memprediksi kelas dengan baik. Adapun secara spesifik untuk nilai *precision* dan *recall* model belum dapat menggambarkan dengan baik pada kategori kelas negara berkembang karena skor yang kecil.
2. Variabel yang signifikan berpengaruh yaitu angka kelahiran dan upah minimum. Angka kelahiran berpengaruh negatif dengan rasio *odds* 0,89 pada negara berkembang dibandingkan dengan tidak maju dan 0,71 pada negara maju dan negara berkembang. Sedangkan upah minimum berpengaruh positif dengan rasio *odds* 11,66 pada negara berkembang dibandingkan dengan tidak maju dan 42,32 pada negara maju dibandingkan dengan negara berkembang yang berarti semakin tinggi angka kelahiran cenderung membawa negara menjadi negara yang tidak maju dan bagi upah minimum semakin tinggi maka akan cenderung membawa negara tersebut diklasifikasikan sebagai negara maju.
3. Variabel berpengaruh tersebut dapat menjadi indikator yang dikhususkan bagi negara yang ingin memperbaiki status klasifikasinya dan fokus memperbaiki yaitu dengan cara mengurangi angka kelahiran dengan program tertentu yang diintervensi pemerintah serta menaikkan secara bertahap upah minimum pada negara tersebut melalui peningkatan kerja pada sektor formal.

DAFTAR PUSTAKA

[1]. Alonso, J. A., Cortez, A. L., & Klasen, S., 2014. LDC and other country groupings: How useful are current approaches to classify countries in a more heterogeneous developing world? *Global Governance and Rules for the Post-2015 Era*, 347.

[2]. Chen, Y. C., 2020. Lecture 7: Multinomial distribution. URL: http://faculty.washington.edu/yenchic/20A_stat512/Lec7_Multinomial.pdf. Diakses pada tanggal 12 Oktober 2023.

[3]. Das, C., Sahoo, A. K., & Pradhan, C., 2022. Multicriteria recommender system using different approaches. In Cognitive Big Data Intelligence with a Metaheuristic Approach. *Academic Press*, 259-277.

[4]. Delnord, M., Hindori- Mohangoo, A. D., Smith, L. K., Szamotulska, K., Richards, J. L., Deb- Rinker, P., ... & Macfarlane, A., 2017. Variations in very preterm birth rates in 30 high- income countries: are valid international comparisons possible using routine data?. *BJOG: An International Journal of Obstetrics & Gynaecology*, *124*(5), 785-794.

[5]. Hosmer DW, Lemeshow S., 2000. *Applied Logistic Regression. Ed ke-2*. New York: John Wiley and Sons.

[6]. Hua, C., Choi, Y. J., & Shi, Q., 2021. Companion to BER 642: Advanced regression methods. *Bookdown. Retrieved August*, *5*, 2021.

[7]. Khan, T., Abimbola, S., Kyobutungi, C., & Pai, M., 2022. How we classify countries and people—and why it matters. *BMJ Global Health*, *7*(6), e009704.

[8]. Li, H., & Zhang, J., 2007. Do High Birth Rates Hamper Economic Growth? *The Review of Economics and Statistics* 2007; 89 (1): 110–117. doi: https://doi.org/10.1162/rest.89.1.110.

[9]. Mirjalili, S. H., & Saadat, H., 2020. How to escape the middle income trap in Iran? Lessons from Malaysia, Thailand South Korea and China. *International Economic Studies*, *50*(1), 1-12.

[10]. Pezzulo, C., Nilsen, K., Carioli, A., Tejedor-Garavito, N., Hanspal, S. E., Hilber, T., James, W. H. M., Ruktanonchai, C. W., Alegana, V., Sorichetta, A., Wigley, A. S. Hornby, G. M., Matthews, Z., & Tatem, A. J., 2021. Geographical distribution of fertility rates in 70 low-income, lower-middle-income, and upper-middle-income countries, 2010–16: a subnational analysis of cross-sectional surveys. *The Lancet Global Health*, *9*(6), e802-e812.

[11]. Suliyanto, 2023. *Buku Ajar Regresi Terapan*. Prodi S-1 Statistika FST Unair, Surabaya.

[12]. Vázquez S. T. dan Sumner A., 2013. Revisiting the Meaning of Development: A Multidimensional Taxonomy of Developing Countries, *The Journal of Development Studies*, 49:12, 1728-1745, DOI: 10.1080/00220388.2013.822071.

[13]. Widiatedja, I. G. N. P., 2021. Indonesia's Export Ban on Nickel Ore: Does It Violate the World Trade Organization (WTO) Rules? *Journal of World Trade*, *55*(4).

[14]. Williams, C. C., & Horodnic, A. V., 2019. Why is informal employment more common in some countries? An exploratory analysis of 112 countries. *Employee Relations: The International Journal*, *41*(6), 1434-1450.