

Faktor Status Bekerja di Kabupaten Karawang 2024 Menggunakan Regresi Logistik Biner Multilevel

Afifa Ulya Rosyida *, Nusar Hajarisman

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

afifaulya16@gmail.com, nusar@unisba.ac.id

Abstract. The multilevel binary logistic regression method is a statistical approach used to analyze the relationship between a binary dependent variable and several independent variables, particularly in datasets with a hierarchical structure, where observation units (such as individuals) are nested within higher-level units (such as regions). In this study, the method is applied to analyze the factors influencing the employment status of the working-age population in Karawang Regency, based on the 2024 National Labor Force Survey (SAKERNAS) data. The dependent variable is employment status, categorized into two groups: employed (1) and unemployed (0). The level-1 independent variables at the individual level include age, gender, highest educational attainment, marital status, and participation in special training. Meanwhile, regional characteristics (urban or rural) are modeled as a level-2 random effect. Parameter estimation is conducted using the Maximum Likelihood method with the Laplace Approximation approach. The analysis results indicate that age, gender, education level, and special training significantly influence employment status. Additionally, there is variation between regions that also affects an individual's likelihood of being employed. These findings highlight the importance of considering regional characteristics when analyzing labor force dynamics in Karawang.

Keywords: *Multilevel Binary Logistic Regression, Maximum Likelihood Estimation.*

Abstrak. Metode regresi logistik biner multilevel merupakan pendekatan statistik yang digunakan untuk menganalisis hubungan antara variabel dependen biner dan sejumlah variabel independen. Dengan konteks data yang memiliki struktur hierarkis, di mana unit pengamatan (seperti individu) dikelompokkan ke dalam unit yang lebih tinggi (seperti wilayah). Dalam penelitian ini, metode tersebut diterapkan untuk menganalisis faktor-faktor yang memengaruhi status bekerja penduduk usia kerja di Kabupaten Karawang berdasarkan data SAKERNAS 2024. Variabel dependen yang diteliti adalah status bekerja dikategorikan menjadi dua: bekerja (1) dan tidak bekerja (0). Variabel independen tingkat individu (level 1) meliputi umur, jenis kelamin, pendidikan terakhir, status perkawinan, dan pelatihan khusus, sedangkan karakteristik wilayah (perkotaan dan perdesaan) dimodelkan sebagai efek acak pada level 2. Penaksiran parameter dilakukan menggunakan metode *Maximum Likelihood* dengan pendekatan *Laplace Approximation*. Hasil analisis menunjukkan bahwa umur, jenis kelamin, pendidikan terakhir, dan pelatihan khusus berpengaruh signifikan terhadap status bekerja. Selain itu, terdapat variasi antar wilayah yang turut memengaruhi kemungkinan seseorang bekerja. Temuan ini menegaskan pentingnya mempertimbangkan karakteristik wilayah dalam analisis ketenagakerjaan di Karawang.

Kata Kunci: *Regresi Logistik Biner Multilevel, Maximum Likelihood Estimation.*

A. Pendahuluan

*afifaulya16@gmail.com

Analisis regresi adalah metode statistik yang digunakan untuk mengidentifikasi pengaruh antara satu variabel dependen dengan satu atau lebih variabel independen (Tampil et al., 2017). Model regresi digunakan tidak hanya untuk memahami hubungan antar variabel, tetapi juga untuk mendeskripsikan data, menaksir parameter, serta melakukan peramalan dan prediksi. Tujuan utamanya adalah melakukan penaksiran parameter dalam model. Analisis regresi telah diterapkan luas dalam berbagai bidang, seperti teknik, ekonomi, manajemen, biologi, dan sosial (Hajarisman, 2010). Dalam perkembangannya, regresi terbagi menjadi regresi linear dan nonlinear. Regresi linear sesuai digunakan ketika variabel dependen bersifat kuantitatif. Namun, jika variabel dependen bersifat kategorik, maka regresi linear tidak tepat, sehingga digunakan pendekatan regresi logistik (Aliyudin & Budyanra, 2016).

Analisis regresi logistik dibagi menjadi beberapa jenis berdasarkan data pada variabel responnya, yaitu regresi logistik biner, multinomial, ordinal, dan poisson. Di antara jenis tersebut, regresi logistik biner merupakan bentuk yang paling sederhana (Saragih et al., 2020). Regresi logistik biner mengasumsikan bahwa setiap pengamatan independen dan memiliki varian yang sama. Namun, asumsi ini tidak terpenuhi pada data bertingkat, seperti individu yang dikelompokkan berdasarkan wilayah (perkotaan/perdesaan), karena pengamatan dalam satu kelompok cenderung lebih mirip dibanding antar kelompok. Untuk itu, digunakan regresi logistik biner multilevel yang memperhitungkan efek acak antar kelompok (Goldstein, 2011).

Regresi logistik biner multilevel merupakan pengembangan dari regresi logistik biner biasa yang digunakan untuk menganalisis data dengan struktur hierarkis. Model ini termasuk dalam kategori *mixed effects model* atau model linear campuran, yang menggabungkan efek tetap (*fixed effect*) dan efek acak (*random effect*) dalam satu kerangka model (Goldstein, 2011). Komponen *fixed effect* mengukur pengaruh variabel bebas yang diasumsikan berlaku secara konsisten di seluruh kelompok, sedangkan *random effect* menangkap variabilitas yang terjadi antar kelompok atau klaster, seperti wilayah tempat tinggal. Model ini dapat memisahkan variasi dari tingkat individu dan kelompok, sehingga estimasinya lebih akurat. Keunggulan utama model ini adalah kemampuannya menghitung *Intraclass Correlation Coefficient* (ICC) yang menunjukkan seberapa besar pengaruh kelompok terhadap outcome individu (Amida & Sitorus, 2021). Dalam proses estimasi parameter biasanya menggunakan *Maximum Likelihood Estimation* (MLE), namun karena struktur data kompleks, digunakan pendekatan numerik seperti *Laplace Approximation* untuk hasil yang lebih optimal (Handayani et al., 2017).

Ketenagakerjaan merupakan salah satu aspek penting dalam pembangunan ekonomi suatu daerah. Dalam perspektif Islam, bekerja merupakan bentuk usaha dalam memenuhi kewajiban individu terhadap diri sendiri, keluarga, dan masyarakat. Tingginya angka penduduk yang bekerja mencerminkan tingginya partisipasi masyarakat dalam aktivitas ekonomi, yang pada akhirnya berkontribusi pada pertumbuhan ekonomi daerah. Kabupaten Karawang dikenal sebagai salah satu kawasan industri utama di Jawa Barat yang menyerap tenaga kerja dari berbagai wilayah. Namun, meskipun potensi ekonominya besar, tantangan ketenagakerjaan masih menjadi isu penting. Berdasarkan data SAKERNAS Tahun 2024, Tingkat Pengangguran Terbuka (TPT) di Karawang tercatat sebesar 8.04%, yang berarti masih terdapat bagian signifikan dari angkatan kerja yang belum terserap dalam dunia kerja meskipun sudah siap bekerja. Isu ini mengindikasikan bahwa keberadaan industri belum sepenuhnya mampu mengatasi masalah pengangguran, kemungkinan akibat ketidaksesuaian antara kebutuhan pasar kerja dan kompetensi tenaga kerja, keterbatasan akses terhadap pelatihan, atau hambatan lain seperti perbedaan wilayah (perkotaan dan perdesaan).

Berdasarkan latar belakang yang telah diuraikan sebelumnya, maka perumusan masalah dalam penelitian ini adalah: “Apa saja faktor-faktor yang memengaruhi status bekerja di Kabupaten Karawang pada tahun 2024 yang dianalisis menggunakan regresi logistik biner multilevel, dan apakah terdapat pengaruh dari karakteristik wilayah, yaitu antara daerah perkotaan dan perdesaan, terhadap status bekerja di wilayah tersebut?” Selanjutnya, tujuan dalam penelitian ini adalah: “Untuk mengetahui faktor-faktor yang memengaruhi status bekerja di Kabupaten Karawang pada tahun 2024 melalui analisis regresi logistik biner multilevel, serta untuk mengetahui apakah terdapat pengaruh karakteristik wilayah perkotaan dan perdesaan terhadap status bekerja di Kabupaten Karawang pada tahun 2024.”

B. Metode

Penelitian ini menggunakan data sekunder yang diperoleh dari hasil Survei Angkatan Kerja Nasional (SAKERNAS) tahun 2024 yang dilakukan oleh Badan Pusat Statistik (BPS). Unit analisis dalam penelitian ini adalah individu usia kerja (≥ 15 tahun) yang berada di Kabupaten Karawang, Provinsi Jawa Barat. Data memiliki struktur hierarkis, di mana individu (level 1) dikelompokkan berdasarkan karakteristik wilayah, yaitu perkotaan atau perdesaan (level 2). Variabel penelitian disajikan pada tabel berikut:

Tabel 1. Variabel Penelitian

Level	Variabel	Kategori	Skala Pengukuran
Individu (Level 1)	Status Bekerja (Y)	0: Tidak Bekerja 1: Bekerja	Nominal
	Umur (X_1)	-	Rasio
	Jenis Kelamin (X_2)	1: Laki-Laki 2: Perempuan	Nominal
	Pendidikan Terakhir (X_3)	1: Tidak/Belum Sekolah 2: SD/MI/SDLB/Paket A 3: SMP/MTs/SMPLB/Paket B 4: SMA/MA/SMLB/Paket C 5: SMK 6: MAK 7: Diploma I/II/III 8: Diploma IV 9: S1 10: S2 11: S2 Terapan 12: S3	Ordinal
	Status Perkawinan (X_4)	1: Belum Kawin 2: Kawin 3: Cerai Hidup 4: Cerai Mati	Nominal
	Pelatihan Khusus (X_5)	1: Pernah 2: Tidak Pernah	Nominal
Wilayah (Level 2)	Karakteristik Wilayah	1: Perkotaan 2: Perdesaan	Nominal

Penelitian ini menggunakan pendekatan model regresi logistik biner multilevel untuk menganalisis faktor-faktor yang memengaruhi status bekerja. Model ini cocok digunakan karena data yang dianalisis memiliki struktur hierarkis, yaitu individu (level-1) yang dikelompokkan dalam wilayah atau klaster (level-2). Berikut adalah tahapan penerapan model regresi logistik biner multilevel:

Membentuk Model Regresi Logistik Biner Multilevel

Model ini digunakan untuk menganalisis data dengan struktur hierarkis multilevel, yaitu individu (level-1) yang dikelompokkan dalam wilayah atau klaster (level-2). Persamaan umum modelnya adalah:

$$\text{logit}(\pi_{ij}) = \beta_{00} + \sum_{q=1}^Q \beta_{0q} Z_{qj} + \sum_{p=1}^p \beta_{p0} X_{p ij} + u_{0j} \quad (1)$$

Keterangan:

π_{ij} = probabilitas individu ke-i pada kelompok ke-j bekerja
 β_{00} = intersep tetap (*fixed intercept*), yaitu rata-rata intersep dari seluruh kelompok
 β_{p0} = koefisien tetap dari variabel prediktor ke-p pada level satu
 X_{pij} = nilai variabel prediktor ke-p untuk individu ke-i dalam kelompok ke-j,
 β_{0q} = koefisien tetap dari variabel prediktor ke-q pada level dua
 Z_{qj} = nilai variabel prediktor ke-q pada level dua untuk kelompok ke-j,
 u_{0j} = efek acak pada level dua, yang diasumsikan terdistribusi normal dengan rata-rata nol dan variansi σ_u^2 .

Menghitung Estimasi

Estimasi parameter β dan σ_u^2 yang diperoleh menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan pendekatan *Laplace Approximation* untuk mengatasi integral kompleks dalam model hierarkis.

Pengujian Signifikansi *Random Effect*

Dalam model regresi logistik dua level, keberadaan struktur level dalam data dapat diketahui melalui analisis menggunakan statistik deviance (Agresti, 2007). Nilai *deviance* berfungsi sebagai indikator yang menunjukkan seberapa baik model memodelkan data observasi. Nilai *deviance* dihitung dari *log-likelihood model*, dan secara umum, semakin kecil nilai *deviance*, maka model dianggap semakin sesuai dengan data.

Hipotesis pada pengujian ini sebagai berikut.

$H_0: \sigma_{u0}^2 = 0 \rightarrow$ artinya tidak ada variasi antar kelompok (efek acak tidak signifikan).

$H_1: \sigma_u^2 \neq 0 \rightarrow$ artinya terdapat variasi antar kelompok (efek acak signifikan).

Persamaan statistik uji *deviance* yang digunakan sebagai berikut.

$$LR = -2 \ln \left[\frac{L_{(0)}}{L_{(r)}} \right] \quad (2)$$

Dimana:

$L_{(0)}$ = nilai *likelihood* dari model logistik tanpa efek acak (tanpa *random effect*),

$L_{(r)}$ = nilai *likelihood* dari model dengan efek acak (dengan *random effect*).

Nilai statistik diatas dibandingkan dengan nilai kritis dari distribusi Chi-kuadrat (χ^2) pada tingkat signifikansi α (5%) dan derajat kebebasan sebesar selisih jumlah parameter antara dua model. H_0 ditolak apabila nilai $LR > \chi_{(\alpha, df)}^2$ atau jika *p-value* $< \alpha$. Penolakan H_0 menunjukkan bahwa efek acak signifikan, sehingga terdapat perbedaan antar kelompok dalam model.

Pengujian signifikansi parameter secara simultan

Pengujian signifikansi secara simultan bertujuan untuk mengevaluasi apakah seluruh variabel prediktor dalam model memiliki pengaruh yang signifikan terhadap variabel respon (Hosmer & Lemeshow, 2000). Hipotesis pada pengujian parameter secara simultan sebagai berikut.

$H_0: \beta_{10} = \beta_{20} = \dots = \beta_{p0} = \beta_{01} = \beta_{0q} = 0$ (secara keseluruhan variabel independen tidak berpengaruh signifikan terhadap variabel dependen)

H_1 : Paling sedikit ada satu $\beta \neq 0$, (paling tidak satu variabel prediktor berpengaruh terhadap variabel dependen)

Persamaan statistik uji G yang digunakan sebagai berikut

$$G^2 = -2 \ln \left[\frac{L_{(null\ model)}}{L_{(conditional\ model)}} \right] \quad (3)$$

Di mana:

$L_{(null\ model)}$ = nilai *likelihood* tanpa variabel prediktor

$L_{(conditional\ model)}$ = nilai *likelihood* dengan variabel prediktor

Statistik G^2 mengikuti distribusi Chi-kuadrat dengan derajat kebebasan sebesar jumlah parameter yang diuji (jumlah total koefisien dari level 1 dan level 2). H_0 ditolak ketika $G^2 > \chi^2_{(\alpha, df)}$ atau $p\text{-value} < \alpha$ ($\alpha = 0,05$). Penolakan H_0 menunjukkan bahwa setidaknya terdapat satu variabel prediktor yang secara simultan berpengaruh signifikan terhadap variabel dependen, sehingga model dengan prediktor lebih tepat digunakan dibandingkan model tanpa prediktor.

Pengujian signifikansi parameter secara parsial

Uji parsial digunakan untuk menilai pengaruh atau signifikansi dari parameter β_{pq} secara individual terhadap variabel respon. Pengujian ini dilakukan menggunakan uji *Wald* (Hosmer & Lemeshow, 2000). Hipotesis pada pengujian parameter secara parsial sebagai berikut.

$H_0: \beta_{pq} = 0$ (variabel prediktor ke-p pada level 1 atau variabel penjelas ke-q pada level 2 secara individu tidak berpengaruh terhadap variabel respon)

$H_1: \beta_{pq} \neq 0$ (variabel prediktor ke-p pada level 1 atau variabel penjelas ke-q pada level 2 secara individu berpengaruh terhadap variabel respon)

Pengujian dilakukan menggunakan statistik uji *Wald* yang dihitung dengan rumus berikut.

$$W = \left[\frac{\hat{\beta}_{pq}}{SE(\hat{\beta}_{pq})} \right] \quad (5)$$

$$SE(\hat{\beta}_{pq}) = \sqrt{Var(\hat{\beta}_{pq})} \quad (6)$$

Dimana:

$\hat{\beta}_{pq}$ = nilai estimasi parameter prediktor ke-p pada level 1 atau variabel penjelas ke-q pada level 2,

$SE(\hat{\beta}_{pq})$ = standar error dari $\hat{\beta}_{pq}$.

Nilai W berdistribusi normal standar. H_0 ditolak jika $|W| > Z_{1-\frac{\alpha}{2}}$ atau $p\text{-value} < \alpha$ (0,05). Penolakan

H_0 menunjukkan bahwa variabel prediktor tersebut berpengaruh signifikan secara parsial terhadap variabel respon.

Odds Ratio

Interpretasi model regresi logistik biner dua level menggunakan *odds ratio*. Nilai odds ratio digunakan untuk menginterpretasi seberapa besar peluang kejadian (bekerja) berubah akibat perubahan satu unit pada variabel prediktor:

$$OR = e^{\hat{\beta}_{pq}} \quad (7)$$

Nilai $OR > 1$ menunjukkan peningkatan peluang bekerja, sedangkan $OR < 1$ menunjukkan penurunan.

Interclass Correlation Coefficient (ICC)

Interclass Correlation Coefficient (ICC) merepresentasikan sejauh mana kesamaan karakteristik dimiliki oleh unit-unit mikro (seperti individu) yang berada dalam kelompok makro yang sama (misalnya wilayah). Nilai ICC dapat dihitung berdasarkan estimasi variansi dari *null model*, dan secara matematis dirumuskan sebagai berikut (Merlo et al., 2006).

$$ICC = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \frac{\pi^2}{3}} \quad (8)$$

Dimana:

σ_{u0}^2 = variansi di level dua (antar kelompok),

$\frac{\pi^2}{3} \approx 3.29$ = variansi di level satu (antar individu).

Nilai varians residual pada level-1 dalam model logistik tetap sebesar $\frac{\pi^2}{3} \approx 3.29$, maka semakin besar nilai ICC menunjukkan bahwa variasi antar kelompok cukup dominan dan model multilevel tepat digunakan.

C. Hasil Penelitian dan Pembahasan

Bagian Hasil Estimasi Parameter

Penelitian ini menerapkan model regresi logistik biner dua level dengan random intercept untuk menganalisis faktor-faktor yang memengaruhi status bekerja individu di Kabupaten Karawang. Estimasi parameter menggunakan metode Maximum Likelihood Estimation (MLE), yang kemudian dilanjutkan dengan pemodelan berdasarkan hasil estimasi dan hasilnya disajikan pada Tabel 2. berikut.

Tabel 2. Estimasi Parameter untuk Model Regresi Logistik Biner Multilevel

Variabel	Koefisien	Standard Error	Nilai Z	p-value
Konstanta ($\hat{\beta}_{00}$)	-0.9810	0.34146	-2.873	0.0041
Umur ($\hat{\beta}_{10}$)	0.0403	0.0036	11.219	<0.0001
Jenis Kelamin ($\hat{\beta}_{20}$)	-0.4087	0.0991	-4.123	<0.0001
Pendidikan Terakhir ($\hat{\beta}_{30}$)	0.1714	0.0176	9.717	<0.0001
Status Perkawinan ($\hat{\beta}_{40}$)	0.0199	0.0444	0.449	0.6536
Pelatihan Khusus ($\hat{\beta}_{50}$)	-0.2425	0.0989	-2.453	0.0142

Untuk memperoleh estimasi pada masing-masing kelompok wilayah, dilakukan perhitungan nilai efek acak (*random effect*). Hasil estimasi menunjukkan adanya perbedaan peluang bekerja antara wilayah perkotaan dan perdesaan. Nilai efek acak disajikan pada Tabel 3. berikut ini.

Tabel 3. Estimasi *Random Effect*

Karakteristik Wilayah	Nilai <i>Random Effect</i>
Perkotaan (u_{01})	0.2443
Perdesaan (u_{02})	-0.2453

Uji Signifikansi Random Effect

Uji signifikansi terhadap *random effect* dilakukan menggunakan *Likelihood Ratio Test* untuk mengetahui apakah terdapat variasi status bekerja antar kelompok wilayah. Berdasarkan hasil pengujian, diperoleh nilai statistik LR sebesar 15.1434 yang melebihi nilai kritis $\chi^2_{0.05;1} = 3.84$, sehingga H_0 ditolak pada taraf signifikansi 5%. Dengan demikian, terdapat *random effect* yang signifikan, yang menunjukkan adanya variasi status bekerja antar wilayah. Oleh karena itu, model regresi logistik biner multilevel dinilai lebih tepat digunakan dalam analisis ini.

Uji Signifikansi Secara Simultan

Uji signifikansi secara simultan dilakukan untuk mengetahui apakah seluruh variabel prediktor secara bersama-sama berpengaruh terhadap variabel respon. Berdasarkan hasil pengujian, diperoleh nilai statistik G^2 sebesar 254.84 dengan p-value < 0.0001 yang melebihi nilai kritis $\chi^2_{0.05;5} = 11.07$, sehingga H_0 ditolak pada taraf signifikansi 5%. Dengan demikian, paling tidak terdapat satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon yaitu status bekerja, sehingga model dengan prediktor dinilai lebih sesuai digunakan dalam analisis ini.

Uji Signifikansi Secara Parsial

Uji signifikansi secara parsial dilakukan menggunakan uji *Wald* untuk mengetahui pengaruh masing-masing variabel prediktor terhadap status bekerja. Berdasarkan hasil pengujian, variabel umur, jenis kelamin, pendidikan terakhir, dan pelatihan khusus memiliki nilai p-value < 0.05 dan nilai *Wald* lebih besar dari $\chi^2_{0.05;1} = 3.841$, sehingga berpengaruh signifikan terhadap status bekerja.

Sementara itu, variabel status perkawinan tidak signifikan karena $p\text{-value} > 0.05$ dan nilai Wald < 3.841 . Dengan demikian, sebagian besar variabel prediktor secara individu berpengaruh terhadap status bekerja di Kabupaten Karawang.

Odds Ratio

Berdasarkan hasil estimasi model regresi logistik biner multilevel, diperoleh bahwa variabel umur (X_1), jenis kelamin (X_2), pendidikan terakhir (X_3), dan pelatihan khusus (X_5) berpengaruh signifikan terhadap status bekerja, sedangkan status perkawinan (X_4) tidak signifikan. Nilai *odds ratio* dari masing-masing variabel signifikan diuraikan sebagai berikut:

1. Umur (X_1)

Koefisien regresi umur sebesar 0.0403 menghasilkan nilai *odds ratio* sebagai berikut:

$$OR_{30 \text{ vs } 20} = e^{0.403} = 1.496$$

Artinya, individu berusia 30 tahun memiliki kemungkinan bekerja sekitar 1.5 kali lebih besar dibandingkan yang berusia 20 tahun.

2. Jenis Kelamin (X_2)

Koefisien regresi jenis kelamin sebesar -0.4087 , *odds ratio* laki-laki terhadap perempuan yaitu

$$OR_{Laki-Laki \text{ vs } Perempuan} = e^{-0.4087} = 0.6646$$

Ini menunjukkan laki-laki memiliki peluang bekerja lebih rendah dibanding perempuan.

3. Pendidikan Terakhir (X_3)

Koefisien pendidikan terakhir sebesar 0.1714 menghasilkan nilai *odds ratio* sebagai berikut:

$$OR_{SMA \text{ vs } SMP} = e^{0.1714} = 1.1869$$

Individu dengan pendidikan SMA memiliki kemungkinan bekerja sekitar 1.19 kali lebih besar dibandingkan lulusan SMP.

4. Pelatihan Khusus (X_5)

Koefisien pelatihan khusus sebesar -0.2425 , menghasilkan nilai *odds ratio* sebagai berikut:

$$OR_{Pernah \text{ Pelatihan vs Tidak Pernah Pelatihan}} = e^{-0.2425} = 0.7847$$

Menunjukkan bahwa individu yang pernah mengikuti pelatihan justru memiliki peluang bekerja sedikit lebih rendah.

5. Efek Acak Wilayah

Perbedaan karakteristik wilayah memengaruhi kemungkinan bekerja. Berdasarkan perhitungan:

$$\frac{\exp[-0.9810 + 0.2443]}{\exp[-0.9810 - 0.2453]} = \exp[0.4895] = 1.63$$

Artinya, individu di wilayah perkotaan memiliki kemungkinan bekerja 1.63 kali lebih besar dibandingkan di perdesaan, dengan asumsi karakteristik individu sama.

Interclass Correlation Coefficient (ICC)

Nilai *interclass correlation coefficient* (ICC) digunakan untuk menggambarkan sejauh mana variasi status bekerja (Y) dipengaruhi oleh perbedaan karakteristik antar wilayah (perkotaan dan perdesaan). Hasil perhitungan ICC menunjukkan nilai sebesar 0.0128. Hal ini berarti bahwa sebesar 1.28% keragaman status bekerja individu di Kabupaten Karawang dapat dijelaskan oleh perbedaan antar wilayah, sedangkan sisanya yaitu sebesar 98.72% berasal dari variasi individu itu sendiri.

Model Terbentuk

Dengan demikian, bentuk model regresi logistik biner dua level dengan *random effect* dapat dirumuskan sebagai berikut.

$$\log\left(\frac{\hat{\pi}_{ij}}{1 - \pi_{ij}}\right) = -0.9810 + 0.0403 \text{ Umur}_{ij} - 0.4087 \text{ Jenis kelamin}_{ij} \\ + 0.1714 \text{ Pendidikan Terakhir}_{ij} \\ + 0.0199 \text{ Status Pernikahan}_{ij} \\ - 0.2425 \text{ Pelatihan Khusus}_{ij} + u_{0j}$$

Sehingga, didapatkan model regresi logistik biner dua level dengan *random effect* untuk setiap kelompok adalah sebagai berikut:

Kelompok 1 (Perkotaan):

$$\log\left(\frac{\hat{\pi}_{ij}}{1 - \pi_{ij}}\right) = (-0.7367) + 0.0403 \text{ Umur}_{ij} - 0.4087 \text{ Jenis kelamin}_{ij} \\ + 0.1714 \text{ Pendidikan Terakhir}_{ij} \\ + 0.0199 \text{ Status Pernikahan}_{ij} \\ - 0.2425 \text{ Pelatihan Khusus}_{ij}$$

Kelompok 2 (Perdesaan):

$$\log\left(\frac{\hat{\pi}_{ij}}{1 - \pi_{ij}}\right) = (-1.2263) + 0.0403 \text{ Umur}_{ij} - 0.4087 \text{ Jenis kelamin}_{ij} \\ + 0.1714 \text{ Pendidikan Terakhir}_{ij} \\ + 0.0199 \text{ Status Pernikahan}_{ij} \\ - 0.2425 \text{ Pelatihan Khusus}_{ij}$$

D. Kesimpulan

Model regresi logistik biner dua level dengan pendekatan *Laplace approximation* berhasil mengidentifikasi faktor-faktor yang memengaruhi status bekerja di Kabupaten Karawang tahun 2024. Hasil analisis menunjukkan bahwa umur, jenis kelamin, pendidikan terakhir, dan pelatihan khusus berpengaruh signifikan terhadap status bekerja. Semakin tinggi usia dan jenjang pendidikan, semakin besar peluang individu untuk bekerja. Perempuan memiliki peluang bekerja lebih tinggi dibanding laki-laki, sedangkan individu yang pernah mengikuti pelatihan justru memiliki peluang bekerja lebih rendah. Status perkawinan tidak berpengaruh signifikan.

Efek karakteristik wilayah juga signifikan, di mana individu yang tinggal di wilayah perkotaan memiliki peluang bekerja 1.63 kali lebih besar dibandingkan perdesaan. Nilai ICC sebesar 1.28% menunjukkan bahwa sebagian kecil variasi status bekerja dapat dijelaskan oleh perbedaan wilayah tempat tinggal. Temuan ini mengindikasikan bahwa meskipun faktor individu dominan, konteks wilayah tetap memberikan kontribusi terhadap status bekerja, sejalan dengan TPT Kabupaten Karawang tahun 2024 yang masih berada di angka 8.04%.

Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada Allah SWT atas rahmat dan hidayah-Nya. Ucapan terima kasih juga disampaikan kepada orang tua dan keluarga atas doa dan dukungannya, serta kepada Bapak Dr. Nusar Hajarisman, M.Si., selaku dosen pembimbing yang telah membimbing penulis. Terima kasih juga kepada semua pihak yang telah membantu dalam proses penelitian ini.

Daftar Pustaka

- Ajeng Mega Pratiwi, & Mutaqin, A. K. (2021). Penerapan Algoritma Naïve Bayes Classifier dalam Memprediksi Status Keberlanjutan Polis Nasabah Asuransi PT.X. *Jurnal Riset Statistika*, 1(2), 117–126. <https://doi.org/10.29313/jrs.v1i2.435>
- Dyar Al Falah Hilman, & Aceng Komarudin Mutaqin. (2023). Penerapan Regresi Double Poisson untuk Memprediksi Pertandingan dan Klasemen Liga 1 Indonesia. *Jurnal Riset Statistika*, 97–106. <https://doi.org/10.29313/jrs.v3i2.2784>
- Fauziah, G., & Sunendiari, S. (2021). Estimasi Pseudo Poisson Maximum Likelihood untuk Mengatasi Masalah dalam Model Log-Linear pada Kasus Kusta di Jawa Barat Tahun 2018. *Jurnal Riset Statistika*, 1(1), 57–62. <https://doi.org/10.29313/jrs.v1i1.147>
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression, Second Edition*.

- Luhung Mustika Budiharti, & Sunendiari, S. (2021). Pemodelan dan Pemetaan Jumlah Penderita Kusta di Jawa Barat dengan Regresi Binomial Negatif dan Flexibly Shaped Spatial Scan Statistic. *Jurnal Riset Statistika*, 1(2), 99–106. <https://doi.org/10.29313/jrs.v1i2.409>
- Zahara, A., & Kurniati, E. (2024). Penerapan Ekstrim Fungsi untuk Menentukan Profit Maksimum dalam Pasar Persaingan Sempurna untuk Short-Run dan Long-Run. *DataMath: Journal of Statistics and Mathematics*, 2(1), 33–40.
- Agresti, A. (2007). An introduction to categorical data analysis (2nd edn). *A John Wiley & Sons, INC., Publication*, 28(11), 1643–1643.
- Aliyudin, F., & Budyanra. (2016). Persalinan Wanita Usia Subur Di Indonesia Menggunakan Data Sdki 2012 (Aplikasi Analisis Regresi Logistik Biner Multilevel). *Jurnal Aplikasi Statistika & Komputasi Statistik*, 8(2086–4132), 33–46.
- Amida, O. V., & Sitorus, J. R. H. (2021). Penerapan Regresi Logistik Biner Multilevel Dalam Analisis Pengaruh Karakteristik Individu, Rumah Tangga, Dan Wilayah Terhadap Status Kemiskinan Balita Di Kepulauan Maluku Dan Pulau Papua. *Seminar Nasional Official Statistics*, 2020(1), 967–977. <https://doi.org/10.34123/semnasoffstat.v2020i1.569>
- Badan Pusat Statistik. (2024). Indikator Pasar Tenaga Kerja Indonesia. In *Badan Pusat Statistik* (Vol. 15, Issue 2, pp. 1–14). <https://www.bps.go.id/id/publication/2024/12/16/d96c70fc6be192f04edf7fe0/indikator-pasar-tenaga-kerja-indonesia-agustus-2024.html>
- Goldstein, H. (2011). *Multilevel Statistical Models (4th ed.)*. Wiley. [https://books.google.co.id/books?id=mdwt7ibSGUYC&lpg=PP13&ots=fAOsFYNW81&dq=Goldstein \(2011\) – Multilevel Statistical Models&lr&hl=id&pg=PP13#v=onepage&q=Goldstein \(2011\) – Multilevel Statistical Models&f=false](https://books.google.co.id/books?id=mdwt7ibSGUYC&lpg=PP13&ots=fAOsFYNW81&dq=Goldstein (2011) – Multilevel Statistical Models&lr&hl=id&pg=PP13#v=onepage&q=Goldstein (2011) – Multilevel Statistical Models&f=false)
- Hajarisman, N. (2010). *Analisis Regresi Lanjut*.
- Handayani, D., Notodiputro, K. A., Sadik, K., & Kurnia, A. (2017). A comparative study of approximation methods for maximum likelihood estimation in generalized linear mixed models (GLMM). *AIP Conference Proceedings*, 1827(December). <https://doi.org/10.1063/1.4979449>
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression, Second Edition*.
- Merlo, J., Chaix, B., Ohlsson, H., Beckman, A., Johnell, K., Hjerpe, P., Råstam, L., & Larsen, K. (2006). A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *Journal of Epidemiology and Community Health*, 60(4), 290–297. <https://doi.org/10.1136/jech.2004.029454>

- Saragih, M. T., Inayah, A. W., Nooraeni, R., Aprilio, M., & Sinsyi, Marita Mutiara Aprilia, Y. R. (2020). Penerapan Regresi Logistik Biner Multilevel pada Partisipasi Angkatan. *Eigen Mathematics Journal*, 3(1), 35–44.
- Tampil, Y., Komaliq, H., & Langi, Y. (2017). Analisis Regresi Logistik Untuk Menentukan Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa FMIPA Universitas Sam Ratulangi Manado. *D'CARTESIAN*, 6(2), 56. <https://doi.org/10.35799/dc.6.2.2017.17023>
- Zakiyatis Salmainsi, & Suliadi. (2021). SPC (Statistical Process Control) Fase II Diagram Kendali Cusum (Cumulative Sum) Nonparametrik Berdasarkan Statistik Mann-Whitney Pada Data Harga Saham PT NO. *Jurnal Riset Statistika*, 1(2), 83–91. <https://doi.org/10.29313/jrs.v1i2.404>