



Klasifikasi status penerima bantuan program keluarga harapan di Provinsi NTB menggunakan metode regresi probit

Lisa Harsyiah^{1*}, Zulhan Widya Baskara¹, Dina Eka Putri¹,
Jurniati²

¹ Statistika, FMIPA, Universitas Mataram, Mataram

⁴ Mahasiswa Statistika, FMIPA, Universitas Mataram, Mataram

lisa_harsyiah@unram.ac.id

Abstract

The Indonesian government's effort to accelerate the achievement of comprehensive social welfare involves adopting strategic policies in the form of distributing social assistance to economically vulnerable communities. One concrete example of this policy is the Family Hope Program (Program Keluarga Harapan/PKH). However, its implementation in the field still faces challenges, particularly in the form of unequal distribution, which has the potential to hinder the program's effectiveness. To address this issue, a rigorous verification system is required to ensure that prospective beneficiaries truly meet the official criteria set by the government. Therefore, classifying households eligible for PKH is a crucial step. The probit regression approach is employed as a method to analyze and determine the household eligibility status. This method yields an accuracy rate of 76.25%, which is considered valid and reliable based on the Press's Q statistic

Keywords: Apparent Error Rate (APER); Classification; PKH; Probit Regression

Abstrak

Upaya Pemerintah Indonesia untuk mempercepat tercapainya kesejahteraan sosial secara menyeluruh yaitu mengambil kebijakan strategis berupa penyaluran bantuan sosial kepada masyarakat yang tergolong rentan secara ekonomi. Salah satu bentuk nyata dari kebijakan ini adalah Program Keluarga Harapan (PKH). Namun, realisasinya di lapangan masih menghadapi hambatan berupa distribusi yang tidak merata, yang berpotensi menghambat efektivitas program. Untuk menanggulangi hal tersebut, diperlukan sistem verifikasi yang ketat guna memastikan bahwa calon penerima manfaat benar-benar memenuhi kriteria resmi dari pemerintah. Oleh karena itu, proses klasifikasi rumah tangga penerima PKH menjadi hal yang krusial. Pendekatan regresi probit digunakan sebagai metode untuk menganalisis dalam menentukan status kepesertaan rumah tangga. Metode ini menghasilkan tingkat akurasi sebesar 76,25%, yang dinilai valid dan andal berdasarkan perhitungan Press's Q.

Kata Kunci: Apparent Error Rate (APER); Klasifikasi; PKH; Regresi Probit

1. PENDAHULUAN

Indonesia merupakan negara berkembang yang selalu mengedepankan kesejahteraan rakyatnya. Upaya Indonesia untuk mempercepat tercapainya kesejahteraan sosial bagi seluruh masyarakat ditempuh melalui penyaluran bantuan sosial kepada pihak-pihak

yang membutuhkan. Salah satu bentuk nyata dari kebijakan tersebut ialah penyelenggaraan Program Keluarga Harapan (PKH) (Sianipar & Cipta, 2023). PKH adalah sebuah bantuan sosial dari pemerintah Indonesia yang memiliki tujuan untuk mendukung percepatan penanganan kemiskinan dan mempercepat terwujudnya kesejahteraan sosial di Indonesia.

Program Keluarga Harapan (PKH) disalurkan secara menyeluruh ke berbagai wilayah di Indonesia. Salah satu provinsi yang teridentifikasi sebagai penerima sekaligus telah merasakan dampaknya ialah masyarakat di Nusa Tenggara Barat (NTB). Menurut Kementerian Sosial Republik Indonesia (2024), jumlah KPM (Keluarga Penerima Manfaat) PKH di NTB mencapai 333.868 KPM, dimana angka ini meningkat dari tahun 2022 yang berjumlah 310.623 KPM (NTB, 2024). Pemerintah berharap bahwa dengan adanya pemberian bantuan PKH ini, kehidupan masyarakat dalam bidang sosial ekonomi dapat lebih sejahtera dan meningkat (Gentara, 2023). Meskipun demikian, peningkatan jumlah KPM PKH tidak menjamin meratanya pembagian PKH kepada masyarakat. Keadaan ini memicu timbulnya berbagai persoalan baru di tengah interaksi antaranggota masyarakat.

Salah satu tantangan utama dalam penyaluran bantuan Program Keluarga Harapan (PKH) adalah ketidaktepatan sasaran, yang kerap memicu ketegangan sosial di tengah Masyarakat (Arimbi, 2022). Untuk mengatasi sekaligus mencegah permasalahan tersebut, diperlukan upaya penegakan kriteria yang sesuai dengan ketentuan resmi yang telah ditetapkan oleh pemerintah. Sejumlah indikator yang dapat dijadikan acuan dalam menetapkan kelayakan penerima bantuan PKH meliputi jumlah anggota keluarga, status kepemilikan rumah tinggal, material atap rumah, material dinding rumah, material lantai rumah, dan sumber air utama.

Berdasarkan kriteria yang telah ditetapkan, langkah penting yang perlu dilakukan adalah mengelompokkan rumah tangga ke dalam kategori penerima dan bukan penerima Program Keluarga Harapan (PKH). Klasifikasi sendiri merupakan pendekatan analitis yang bertujuan mengidentifikasi pola diskriminatif dalam data guna menentukan kategori suatu entitas berdasarkan karakteristik dan perilaku kelompok yang telah terdefinisi sebelumnya (Romli & Zy, 2020). Inti dari proses ini adalah untuk memetakan data baru ke dalam kelas yang telah memiliki label yang pasti (Johnson & Wichern, 2007). Selain berfungsi sebagai alat kategorisasi, klasifikasi juga berperan dalam menilai performa suatu model (Dewanti, 2019). Beragam teknik dapat digunakan dalam proses ini, salah satunya adalah pendekatan regresi probit yang terbukti efektif dalam konteks serupa (Epriliyanti & Ratnasari, 2020).

Regresi probit adalah jenis regresi dengan tujuan untuk menentukan keterkaitan diantara variabel independen terhadap variable dependen (Ariessela, 2021). Model regresi probit merepresentasikan bentuk fungsi distribusi kumulatif yang berasal dari

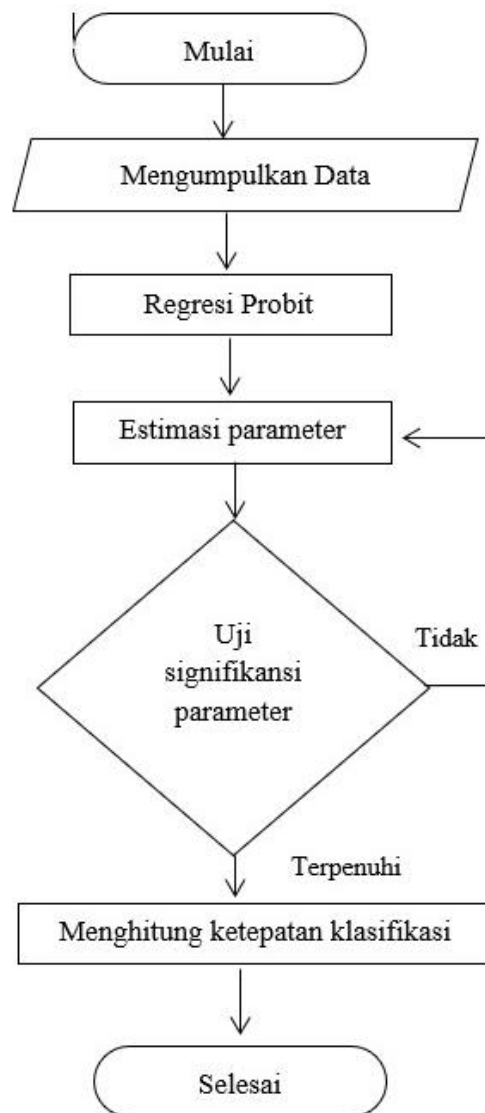
distribusi normal (Fathurahman, 2019). Kelebihan regresi probit adalah mampu menangani data biner dengan baik dan dapat diinterpretasikan dengan mudah karena hasilnya dinyatakan dalam probabilitas. Penelitian mengenai regresi probit telah dilakukan oleh (Febyanti, 2022) untuk memodelkan faktor-faktor yang berpengaruh terhadap harga rumah di Jabodetabek. Selanjutnya penelitian yang dilakukan oleh (Fatmallah & Baskara, 2021) dimana digunakan regresi probit biner untuk menganalisis faktor-faktor yang memengaruhi Indeks Prestasi Kumulatif mahasiswa Matematika FMIPA Universitas Mataram. Adapun hasil penelitian menunjukkan tingkat akurasi klasifikasi yang dicapai adalah sebesar 68,12%. Penelitian lainnya mengenai PKH yang dilakukan oleh (Harsyiah & Hadijati, 2024) yang membahas mengenai klasifikasi kelayakan status penerima bantuan Program Keluarga Harapan (PKH) menggunakan analisis diskriminan dan *naive bayes*. Hasil penelitian menunjukkan bahwa analisis diskriminan memperoleh ketepatan klasifikasi lebih baik dibandingkan naive bayes yaitu sebesar 75.5 %. Penelitian lainnya mengenai PKH juga dilakukan oleh (Manik, 2020) yang membahas tentang penerapan MOORA dalam mendukung keputusan kelayakan penerima bantuan Program Keluarga Harapan.

Berdasarkan penjelasan tersebut, perlu dilakukan penelitian untuk mengklasifikasikan status penerima bantuan sosial PKH di NTB menggunakan regresi probit sehingga diperoleh model klasifikasi yang memiliki ketepatan klasifikasi lebih tinggi dibandingkan penelitian sebelumnya serta faktor apa saja yang mempengaruhi status kelayakan penerima bantuan PKH. Hasil penelitian diharapkan dapat memberikan masukan terhadap pihak terkait sehingga menjadi pertimbangan apakah suatu rumah tangga layak atau tidak dalam mendapatkan bantuan PKH.

2. METODE PENELITIAN

Regresi probit digunakan sebagai metode untuk menganalisis data sekunder yaitu seluruh masyarakat di 10 Kota/ Kabupaten Penerima bantuan PKH di Provinsi Nusa Tenggara Barat yang diperoleh dari hasil Survei Sosial dan Ekonomi Nasional (SUSENAS) tahun 2022 yang bisa diakses melalui <https://silastik.bps.go.id>. Penelitian ini memfokuskan analisis pada klasifikasi penerima Program Keluarga Harapan (PKH) dengan faktor-faktor yang mempengaruhi seperti jumlah anggota keluarga (X_1), status kepemilikan rumah tinggal (X_2), material atap rumah (X_3), material dinding rumah (X_4), jenis lantai rumah (X_5), sumber air utama (X_6), serta kepemilikan barang seperti kulkas (X_7), komputer/laptop (X_8), emas (X_9), motor (X_{10}), mobil (X_{11}), dan televisi (X_{12}).

Langkah-langkah dalam melakukan analisis data menggunakan regresi probit dapat dilihat pada diagram alir pada Gambar 1 berikut.



Gambar 1. Diagram Alir Penelitian

Adapun penjelasan langkah-langkah penelitian pada **Gambar 1** adalah (Nur'eni dan Handayani, 2020):

1. Melakukan pengujian normal multivariat. Asumsi awal yang harus dipenuhi dalam pengklasifikasian menggunakan metode Regresi Probit adalah data berdistribusi normal multivariat. Pengujian kenormalan dapat dilakukan dengan menghitung nilai jarak kuadrat setiap observasi menggunakan rumus berikut:

$$d_i^2 = (X_i - \bar{X})' S^{-1} (X_i - \bar{X}) \quad (1)$$

2. Melakukan estimasi parameter. Metode penaksiran parameter yang digunakan regresi probit adalah *Maximum Likelihood Estimator* (MLE).
3. Melakukan uji signifikansi parameter

Uji signifikansi parameter dapat dibagi menjadi dua yaitu uji signifikansi secara parsial dengan menggunakan uji Wald dan uji signifikansi secara simultan dengan menggunakan uji G. Adapun persamaan statistic uji Wald adalah sebagai berikut:

$$W_j = \frac{\widehat{\beta}_j}{SE(\widehat{\beta}_j)} \quad (2)$$

Dan persamaan statistic uji G adalah sebagai berikut:

$$G = -2 \ln \left[\frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_0}{n}\right)^{n_0}}{\prod_{i=1}^n P_i^{y_i} (1-P_i)^{1-y_i}} \right] \quad (3)$$

4. Melakukan uji kesesuaian model. Menguji kesesuaian model regresi probit dapat dilakukan dengan uji Deviance dengan persamaan statistik uji sebagai berikut:

$$D = -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{P_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1-P_i}{1-y_i} \right) \right] \quad (4)$$

5. Melakukan perhitungan akurasi klasifikasi. Keakuratan klasifikasi atau pengelompokan dapat dilakukan dengan menghitung Apparent Error Rate (APER), Hit Rasio, dan Press's Q.

3. HASIL DAN PEMBAHASAN

3.1 Uji normal multivariat

Perhitungan pengujian normalitas untuk d_1^2 (pengamatan 1) menggunakan persamaan (1) adalah sebagai berikut :

$$\begin{aligned} d_1^2 &= (X_i - \bar{X})^T S^{-1} (X_i - \bar{X}) \\ &= \begin{bmatrix} 1 - 0.8200 \\ 3 - 2.7125 \\ \vdots \\ 0 - 0.1325 \end{bmatrix}^T \begin{bmatrix} 7.2914 & 0.0293 & \dots & -0.2766 \\ 0.0293 & 1.8510 & \dots & -0.1191 \\ \vdots & \vdots & \ddots & \vdots \\ -0.2766 & -0.1191 & \dots & 11.9545 \end{bmatrix} \begin{bmatrix} 1 - 0.8200 \\ 3 - 2.7125 \\ \vdots \\ 0 - 0.1325 \end{bmatrix} \\ &= 5.6684 \end{aligned}$$

Langkah yang sama dilakukan untuk perhitungan nilai d_2^2 sampai dengan d_{400}^2 .

Nilai q_i dihitung dengan rumus berikut :

$$q_i = \left(\frac{i - 0.5}{n} \right)$$

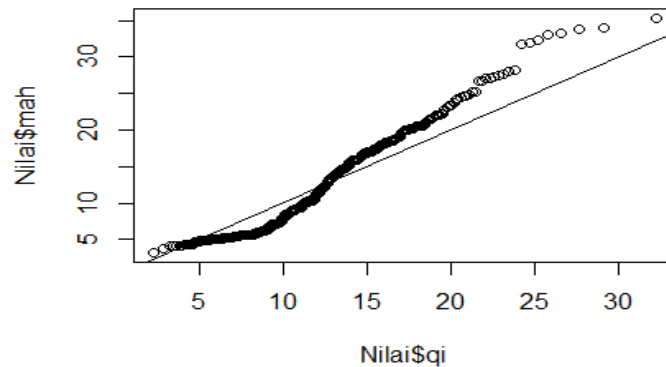
Perhitungan untuk q_1 sebagai berikut :

$$\begin{aligned} q_1 &= \left(\frac{1 - 0.5}{400} \right) \\ &= 0.00125 \end{aligned}$$

Langkah yang sama dilakukan untuk perhitungan nilai q_2 sampai dengan q_{400} .

Jarak kuadrat masing-masing pengamatan yang telah diperoleh disusun secara berurutan mulai dari nilai terkecil hingga terbesar. Setelah itu, dilakukan visualisasi

dengan memetakan nilai d_i^2 terhadap q_i . Gambar 2 berikut merupakan penyajian hasil pengujian normalitas data.



Gambar 2. Uji Normal Multivariat Data

Pada **Gambar 2** terlihat titik-titik pada plot tersusun sedemikian rupa sehingga membentuk pola linier yang jelas, menandakan bahwa distribusi data mendekati normal multivariat.

3.2 Estimasi Parameter

Regresi probit biner menggunakan prinsip estimasi *Maximum Likelihood*. Karena hasil penaksiran parameter yang diperoleh dengan metode *Maximum Likelihood* merupakan fungsi implisit, maka metode iterasi Newton Rhaspon digunakan untuk menaksir parameter. Berikut adalah perhitungan estimasi parameter probit biner dengan menggunakan metode Newton Rhaspon:

1. Menghitung nilai taksiran awal parameter β^0 menggunakan OLS yaitu:

$$\beta^0 = [X^T X]^{-1} X^T Y$$

$$= \left(\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{1,1} & x_{2,1} & \dots & x_{400,1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,12} & x_{2,12} & \dots & x_{400,12} \end{bmatrix} \begin{bmatrix} 1 & x_{1,1} & \dots & x_{1,12} \\ 1 & x_{2,1} & \dots & x_{2,12} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{400,1} & \dots & x_{400,12} \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{1,1} & x_{2,1} & \dots & x_{400,1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,12} & x_{2,12} & \dots & x_{400,12} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{400} \end{bmatrix}$$

$$\beta^0 = \begin{bmatrix} 0.41894 \\ 0.16274 \\ 0.07197 \\ -0.01173 \\ 0.03779 \\ 0.02946 \\ 0.04030 \\ 0.10386 \\ 0.07906 \\ -0.01471 \\ -0.02729 \\ 0.06863 \\ 0.10512 \end{bmatrix}$$

2. Membentuk matriks turunan pertama $g^{(0)}\beta^{(0)}$

Adapun nilai turunan pertama dari fungsi *likelihood* terhadap β_0 sampai β_{12} adalah sebagai berikut:

$$g^{(0)}\beta^{(0)} = \begin{bmatrix} -12.68742 \\ -2.48010 \\ -59.60980 \\ -39.85673 \\ -19.12741 \\ 2.55385 \\ 27.63124 \\ 13.37532 \\ 10.49473 \\ 10.05115 \\ -4.31963 \\ 9.34244 \\ 13.96876 \end{bmatrix}$$

3. Membentuk matriks turunan kedua $H^{(0)}\beta^{(0)}$

$$\frac{\partial^2 \ln(\beta)}{\partial \beta_0^2}$$

Langkah yang sama untuk menghitung nilai turunan kedua dari fungsi *likelihood* terhadap $\beta_0\beta_2$ sampai $\beta_{11}\beta_{12}$.

Setelah dilakukan perhitungan, maka matriks hessian yang terbentuk adalah :

$$H^{(0)}\beta^{(0)} = \begin{bmatrix} 206.00111 & 167.06962 & \cdots & 23.53203 \\ 167.06962 & 167.06962 & \cdots & 17.82419 \\ \vdots & \vdots & \ddots & \vdots \\ 23.53203 & 17.82419 & \cdots & 23.53203 \end{bmatrix}$$

Invers matriks $H^{(0)}\beta^{(0)}$ adalah:

$$(H^{(0)}\beta^{(0)})^{-1} = \begin{bmatrix} 0.28111 & -0.03398 & \cdots & 0.02246 \\ -0.03398 & 0.03444 & \cdots & -0.00148 \\ \vdots & \vdots & \ddots & \vdots \\ 0.02246 & -0.00148 & \cdots & 0.06397 \end{bmatrix}$$

4. Menghitung perkalian matriks $(H^{(0)}\beta^{(0)})^{-1}$ dan matriks $g^{(0)}\beta^{(0)}$

$$(H^{(0)}\beta^{(0)})^{-1} g^{(0)}\beta^{(0)} = \begin{bmatrix} 0.28111 & -0.03398 & \cdots & 0.02246 \\ -0.03398 & 0.03444 & \cdots & -0.00148 \\ \vdots & \vdots & \ddots & \vdots \\ 0.02246 & -0.00148 & \cdots & 0.06397 \end{bmatrix} \begin{bmatrix} -12.68742 \\ -2.48010 \\ \vdots \\ 13.96876 \end{bmatrix}$$

$$= \begin{bmatrix} -0.66789 \\ 0.37175 \\ 0.20167 \\ -0.02373 \\ 0.07690 \\ 0.06315 \\ 0.09302 \\ 0.21230 \\ 0.24042 \\ -0.00771 \\ -0.05802 \\ 0.22608 \\ 0.33440 \end{bmatrix}$$

5. Proses iterasi pertama dilakukan dengan persamaan berikut :

$$\beta^{(m+1)} = \beta^{(m)} + (H^{(m)}\beta^{(m)})^{-1}g^{(m)}\beta^{(m)}$$

$$\beta^{(1)} = \beta^{(0)} + (H^{(0)}\beta^{(0)})^{-1}g^{(0)}\beta^{(0)}$$

$$\beta^{(1)} = \begin{bmatrix} 0.41894 \\ 0.16274 \\ 0.07197 \\ -0.01173 \\ 0.03779 \\ 0.02946 \\ 0.04030 \\ 0.10386 \\ 0.07906 \\ -0.01471 \\ -0.02729 \\ 0.06863 \\ 0.10512 \end{bmatrix} + \begin{bmatrix} -0.66789 \\ 0.37175 \\ 0.20167 \\ -0.02373 \\ 0.07690 \\ 0.06315 \\ 0.09302 \\ 0.21230 \\ 0.24042 \\ -0.00771 \\ -0.05802 \\ 0.22608 \\ 0.33440 \end{bmatrix} = \begin{bmatrix} -0.24895 \\ 0.53449 \\ 0.27364 \\ -0.03546 \\ 0.11370 \\ 0.09261 \\ 0.13332 \\ 0.31616 \\ 0.31948 \\ -0.02242 \\ -0.08531 \\ 0.29471 \\ 0.43952 \end{bmatrix}$$

Selanjutnya nilai $\beta^{(1)}$ digunakan sebagai nilai taksiran awal untuk iterasi kedua. Langkah yang sama dilakukan untuk melakukan perhitungan iterasi kedua dan seterusnya sampai proses iterasi mendapatkan penaksir parameter yang konvergen. Proses iterasi berakhir ketika $|\beta^{(m-1)} - \beta^{(m)}| < \varepsilon$. Berdasarkan hasil perhitungan diperoleh nilai estimasi parameter pada iterasi $m = 7$ yang ditampilkan pada **Tabel 1** sebagai berikut.

Tabel 1. Estimasi parameter regresi probit

Parameter	Estimasi
β_0	-0.02428
β_1	0.63725
β_2	0.41602
β_3	-0.02211
β_4	0.11387
β_5	0.09281
β_6	0.15658

Parameter	Estimasi
β_7	0.27672
β_8	0.59625
β_9	0.7121
β_{10}	-0.08209
β_{11}	0.69366
β_{12}	1.02526

Diperoleh model regresi probit biner sebagai hasil dari estimasi parameter yang telah dilakukan, yaitu sebagai berikut:

$$Y^* = -0.02428 + 0.63725x_1 + 0.41602x_2 - 0.02211x_3 + 0.11387x_4 + 0.09281x_5 + 0.15658x_6 \\ + 0.27672x_7 + 0.59625x_8 + 0.07121x_9 - 0.08209x_{10} + 0.69366x_{11} \\ + 1.02526x_{12}$$

3.3. Uji Signifikansi Parameter Secara Parsial

Pengujian signifikansi terhadap parameter secara parsial dapat menggunakan hipotesis uji Wald sebagai berikut :

$$H_0: \beta_j = 0$$

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

Statistik uji Wald yang digunakan adalah:

$$W_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

Dimana,

$$SE(\beta_j) = \sqrt{\text{var}(\beta_j)}$$

$$\text{var}(\beta_j) = [X^T \hat{V} X]^{-1}$$

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \dots & 0 \end{bmatrix} \begin{bmatrix} 0.63648 & 0 & \dots & 0 \\ 0 & 0.60809 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0.52011 \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{bmatrix}^{-1} \\ = \begin{bmatrix} 1.72024 & 0.01735 & \dots & -0.02095 \\ 0.01735 & 0.02835 & \dots & -0.00032 \\ \vdots & \vdots & \ddots & \vdots \\ -0.02095 & -0.00032 & \dots & 0.02693 \end{bmatrix}$$

Berdasarkan perhitungan matriks $[X^T \hat{V} X]^{-1}$ diambil nilai diagonalnya untuk menghitung nilai SE sebagai berikut.

$$\begin{bmatrix} \text{var}\beta_0 \\ \text{var}\beta_1 \\ \text{var}\beta_2 \\ \text{var}\beta_3 \\ \text{var}\beta_4 \\ \text{var}\beta_5 \\ \text{var}\beta_6 \\ \text{var}\beta_7 \\ \text{var}\beta_8 \\ \text{var}\beta_9 \\ \text{var}\beta_{10} \\ \text{var}\beta_{11} \\ \text{var}\beta_{12} \end{bmatrix} = \begin{bmatrix} 0.36474 \\ 0.04101 \\ 0.02091 \\ 0.00626 \\ 0.01880 \\ 0.00462 \\ 0.00443 \\ 0.02778 \\ 0.10998 \\ 0.07463 \\ 0.02994 \\ 0.24003 \\ 0.21751 \end{bmatrix}$$

Sehingga diperoleh nilai W_j sebagai berikut:

$$\begin{aligned} W_0 &= \frac{\widehat{\beta}_0}{\sqrt{\text{var}(\widehat{\beta}_0)}} \\ &= \frac{-0.02428}{0.60393} \\ &= -0.040 \end{aligned}$$

Langkah yang sama dilakukan untuk perhitungan nilai W_1 sampai dengan W_{12} dengan nilai Z_{tabel} sebagai berikut:

$$\begin{aligned} Z_{tabel} &= Z_{\alpha/2} \\ &= Z_{0.025} \\ &= 1.96 \end{aligned}$$

Hasil perhitungan uji Wald disajikan pada Tabel 2 sebagai berikut.

Tabel 2. Nilai uji parsial regresi probit

Parameter	Estimasi	$SE(\beta_j)$	$ W_j $	$P\text{-value}$	Kesimpulan
β_0	-0.02428	0.60393	-0.040	0.96793	Tidak signifikan
β_1	0.63725	0.20250	3.147	0.00165	Signifikan
β_2	0.41602	0.14459	2.877	0.00401	Signifikan
β_3	-0.02211	0.07913	-0.279	0.77995	Tidak signifikan
β_4	0.11387	0.13711	0.831	0.40623	Tidak signifikan
β_5	0.09281	0.06794	1.366	0.17193	Tidak signifikan

Parameter	Estimasi	SE(β_j)	$ W_j $	P-value	Kesimpulan
β_6	0.15658	0.06654	2.353	0.01862	Signifikan
β_7	0.27672	0.16666	1.660	0.09685	Tidak signifikan
β_8	0.59625	0.33163	1.798	0.07219	Tidak signifikan
β_9	0.07121	0.27318	0.261	0.79435	Tidak signifikan
β_{10}	-0.08209	0.17303	-0.474	0.63519	Tidak signifikan
β_{11}	0.69366	0.48992	1.416	0.15682	Tidak signifikan
β_{12}	1.02526	0.46637	2.198	0.02792	Signifikan

Berdasarkan **Tabel 2** diperoleh 3 variabel yang signifikan karena nilai uji Wald $|W_j| > 1.96$ atau $p\text{-value} < 0.05$. Variabel yang berpengaruh signifikan adalah variabel jumlah anggota keluarga, kepemilikan rumah tinggal, sumber air utama dan kepemilikan televisi.

3.4 Uji Signifikansi Parameter Secara Simultan

Pengujian signifikansi parameter secara simultan dilakukan dengan uji G atau rasio likelihood dengan hipotesis di bawah ini.

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

$$H_1: \text{paling tidak terdapat satu } \beta_j \neq 0 ; j = 1, 2, \dots, p$$

Statistik uji G :

$$G = -2 \ln \left[\frac{\binom{n_1}{n}^{n_1} \binom{n_0}{n}^{n_0}}{\prod_{i=1}^n P_i^{y_i} (1-P_i)^{1-y_i}} \right]$$

$$\begin{aligned} G &= 2 \left\{ \sum_{i=1}^n [y_i \ln (F(x_i \beta)) + (1 - y_i) \ln (1 - F(x_i \beta))] - [n_0 \ln(n_0) + n_1 \ln(n_1) - n \ln(n)] \right\} \\ &= 2[(-0.30516) + (-0.55283) + \dots + (-0.38622)] - [101 \ln(101) + 299 \ln(299) - 400 \ln(400)] \\ &= 2[(-188.25235) - (-226.02602)] \\ &= 75.54733 \end{aligned}$$

Perhitungan untuk nilai χ^2_{tabel} sebagai berikut :

$$\chi^2_{tabel} = \chi^2_{\alpha, v}$$

$$\begin{aligned}
 &= \chi^2_{0.05,12} \\
 &= 21.02607
 \end{aligned}$$

Hasil uji G disajikan pada **Tabel 3** berikut.

Tabel 3. Nilai uji simultan regresi probit

<i>Likelihood</i>	<i>likelihoodNull</i>	G
-188.25235	-2226.02602	75.54733

Berdasarkan **Tabel 3** nilai G yang diperoleh sebesar 75.54733 dan nilai χ^2_{tabel} sebesar 21.02607. Hal ini mengartikan bahwa $G > \chi^2_{tabel}$ maka H_0 ditolak sehingga dapat disimpulkan bahwa keseluruhan variabel independen secara bersama-sama memberikan kontribusi signifikan terhadap penentuan status penerima bantuan sosial Program Keluarga Harapan di wilayah Nusa Tenggara Barat.

3.5 Uji Kesesuaian Model

Perbedaan antara hasil observasi dengan hasil prediksi dapat diketahui dengan menguji kesesuaian model terhadap hipotesis sebagai berikut:

H_0 : Model sesuai

H_1 : Model tidak sesuai

Statistik uji yang digunakan adalah uji *Deviance*:

$$\begin{aligned}
 D &= -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{p_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1-p_i}{1-y_i} \right) \right] \\
 &= -2 [(-0.30516) + (-0.55283) + \dots + (-0.38622)] \\
 &= 376.50471
 \end{aligned}$$

Tabel 4. Nilai Uji *Deviance*

<i>Deviance</i>	Derajat Bebas
376.50471	387

Berdasarkan **Tabel 4** diperoleh nilai *Deviance* sebesar 376.50471 dan diperoleh nilai χ^2_{tabel} sebesar 433.86974 ini mengartikan bahwa nilai $D < \chi^2_{tabel}$ maka tidak ada alasan untuk menolak H_0 sehingga dapat diambil kesimpulan bahwa model sesuai.

3.6 Menginterpretasikan Model

Langkah selanjutnya yang harus dilakukan dalam regresi probit adalah menghitung nilai probabilitas dan nilai efek marginal pada setiap pengamatan dengan menggunakan variabel independen yang signifikan. Variabel yang signifikan terhadap pengelompokan mengenai identifikasi penerima bantuan PKH di Nusa Tenggara Barat adalah variabel jumlah anggota keluarga (x_1), status kepemilikan rumah tinggal (x_2), sumber air utama (x_6), dan kepemilikan televisi (x_{12}).

Probabilitas model probit biner untuk $Y = 0$ adalah :

$$P(Y = 0) = 1 - \Phi(0.63725x_1 - 0.41602x_2 + 0.15658x_6 + 1.02526x_{12})$$

Probabilitas model probit biner untuk $Y = 1$ adalah:

$$P(Y = 1) = \Phi(0.63725x_1 + 0.41602x_2 + 0.15658x_6 + 1.02526x_{12})$$

Untuk pengamatan pertama dengan nilai $x_1 = 1$, $x_2 = 3$, $x_6 = 2$ dan $x_{12} = 0$ maka fungsi probabilitas yang diperoleh sebagai berikut:

$$\begin{aligned} P(Y = 0) &= 1 - \Phi(0.63725(1) + 0.41602(3) + 0.15658(2) + 1.02526(0)) \\ &= 1 - \Phi(2.19847) \\ &= 0.0140 \end{aligned}$$

$$\begin{aligned} P(Y = 1) &= \Phi(0.63725(1) + 0.41602(3) + 0.15658(2) + 1.02526(0)) \\ &= \Phi(2.19847) \\ &= 0.9860 \end{aligned}$$

Berdasarkan perhitungan diatas untuk pengamatan pertama diperoleh probabilitas untuk menerima bantuan sebesar 0.0140 atau 1.40% sedangkan probabilitas untuk tidak menerima bantuan sebesar 0.9860 atau 98.60%. Selanjutnya dilakukan perhitungan yang sama untuk pengamatan lainnya.

Efek marginal dengan asumsi variabel lain tetap dapat digunakan untuk menentukan besar kecilnya pengaruh perubahan variabel independen terhadap variabel dependen. Efek marginal dari variabel independen yang signifikan untuk pengamatan pertama yaitu dengan mencari turunan dari fungsi peluang setiap kategori (Y), yaitu:

$$\begin{aligned} \frac{\partial(Y = 0)}{\partial X_1} &= -\phi(X^T \beta) \beta_1 \\ &= -(0.63725) \phi \left(\begin{array}{l} 0.63725(1) + 0.41602(3) \\ +0.15658(2) + 1.02526(0) \end{array} \right) \\ &= -0.63725 \phi(2.19847) \\ &= -0.63725 (0.03559) \\ &= -0.0227 \end{aligned}$$

$$\begin{aligned} \frac{\partial(Y = 1)}{\partial X_1} &= \phi(X^T \beta) \beta_1 \\ &= 0.63725 \phi \left(\begin{array}{l} 0.63725(1) + 0.41602(3) \\ +0.15658(2) + 1.02526(0) \end{array} \right) \\ &= 0.63725 \phi(2.19847) \\ &= 0.63725 (0.03559) \\ &= 0.0227 \end{aligned}$$

Berdasarkan hasil perhitungan yang telah dilakukan, dapat diketahui nilai efek marginal dari variabel jumlah anggota keluarga (x_1) untuk pengamatan pertama. Nilai efek marginal yang diperoleh sebesar -0.0227 memberikan arti bahwa jumlah anggota rumah tangga yang lebih dari 4 orang akan menurunkan probabilitas pengamatan pertama sebesar 2.27% untuk menerima bantuan. Sedangkan nilai efek marginal yang diperoleh sebesar 0.0227 memberikan arti bahwa jumlah anggota rumah tangga yang lebih dari 4 orang akan menaikkan probabilitas pengamatan pertama sebesar 2.27% untuk tidak menerima bantuan.

$$\frac{\partial(Y = 0)}{\partial X_2} = -\phi(X^T \beta) \beta_2$$

$$\begin{aligned}
&= -0.41602 \phi \left(\begin{array}{l} 0.63725(1) + 0.41602(3) \\ +0.15658(2) + 1.02526(0) \end{array} \right) \\
&= -0.41602 \phi(2.19847) \\
&= -0.41602 (0.03559) \\
&= -0.0148 \\
\frac{\partial(Y = 0)}{\partial X_2} &= \phi(X^T \beta) \beta_2 \\
&= 0.41602 \phi \left(\begin{array}{l} 0.63725(1) + 0.41602(3) \\ +0.15658(2) + 1.02526(0) \end{array} \right) \\
&= 0.41602 \phi(2.19847) \\
&= 0.41602 (0.03559) \\
&= 0.0148
\end{aligned}$$

Hasil perhitungan menunjukkan nilai efek marginal dari status kepemilikan rumah tinggal (x_2) sebesar -0.0148 memberikan arti bahwa rumah tangga yang memiliki rumah milik sendiri akan menurunkan probabilitas pengamatan pertama sebesar 1.48% untuk menjadi penerima bantuan. Sedangkan nilai efek marginal sebesar 0.0148 memberikan arti bahwa rumah tangga yang memiliki rumah milik sendiri akan menaikkan probabilitas pengamatan pertama sebesar 1.48% untuk tidak menerima bantuan.

$$\begin{aligned}
\frac{\partial(Y = 0)}{\partial X_6} &= -\phi(X^T \beta) \beta_6 \\
&= -0.15658 \phi \left(\begin{array}{l} 0.63725(1) + 0.41602(3) \\ +0.15658(2) + 1.02526(0) \end{array} \right) \\
&= -0.15658 \phi(2.19847) \\
&= -0.15658 (0.03559) \\
&= -0.0056
\end{aligned}$$

$$\begin{aligned}
\frac{\partial(Y = 1)}{\partial X_6} &= \phi(X^T \beta) \beta_6 \\
&= 0.15658 \phi \left(\begin{array}{l} 0.63725(1) + 0.41602(3) \\ +0.15658(2) + 1.02526(0) \end{array} \right) \\
&= 0.15658 \phi(2.19847) \\
&= 0.15658 (0.03559) \\
&= 0.0056
\end{aligned}$$

Perhitungan yang diperoleh menunjukkan bahwa efek marginal dari variabel sumber air utama (x_6) untuk pengamatan pertama sebesar -0.0056 memberikan arti bahwa sumber air utama yang diperoleh dari sumur/mata air terlindung akan menurunkan probabilitas pengamatan pertama sebesar 0.56% untuk menerima bantuan. Sedangkan nilai efek marginal yang diperoleh sebesar 0.0056 memberikan arti bahwa sumber air utama yang diperoleh dari sumur bor/pompa akan menaikkan probabilitas pengamatan pertama sebesar 0.56% untuk tidak menerima bantuan.

$$\begin{aligned}
\frac{\partial(Y = 0)}{\partial X_{12}} &= -\phi(X^T \beta) \beta_{12} \\
&= -1.02526 \phi \left(\begin{array}{l} 0.63725(1) + 0.41602(3) \\ +0.15658(2) + 1.02526(0) \end{array} \right)
\end{aligned}$$

$$\begin{aligned}
&= -1.02526 \phi(2.19847) \\
&= -1.02526 (0.03559) \\
&= -0.0365 \\
\frac{\partial(Y = 1)}{\partial X_{12}} &= \phi(X^T \beta) \beta_{12} \\
&= 1.02526 \phi \left(\begin{array}{l} 0.63725(1) + 0.41602(3) \\ +0.15658(2) + 1.02526(0) \end{array} \right) \\
&= 1.02526 \phi(2.19847) \\
&= 1.02526 (0.03559) \\
&= 0.0365
\end{aligned}$$

Perhitungan yang diperoleh menunjukkan bahwa efek marginal dari variabel kepemilikan televisi (x_6) untuk pengamatan pertama sebesar -0.0365 memberikan arti bahwa rumah tangga yang tidak memiliki televisi akan menurunkan probabilitas pengamatan pertama sebesar 3.65% untuk menerima bantuan. Sedangkan nilai efek marginal yang diperoleh sebesar 0.0365 memberikan arti bahwa rumah tangga yang tidak memiliki televisi akan menaikkan probabilitas pengamatan pertama sebesar 3.65% untuk tidak menerima bantuan.

Berdasarkan perhitungan efek marginal diperoleh hasil yang dapat memberikan masukan bagi pihak terkait bahwa banyaknya anggota keluarga (lebih dari 4), rumah tangga yang belum memiliki rumah tinggal, rumah tangga yang belum memiliki akses air bersih serta belum memiliki televisi lebih dipertimbangkan sebagai penerima bantuan PKH.

3.7 Menentukan Ketepatan Klasifikasi

Berdasarkan model regresi probit yang didapatkan maka dibentuk tabel klasifikasi seperti pada Tabel 5 sebagai berikut:

Tabel 5. Klasifikasi regresi probit

<i>Actual Group</i>	<i>Predicted Group</i>	
	0	1
0	17	84
1	11	288

Selanjutnya dilakukan perhitungan APER, *Hit Rasio* dan *Press's Q* berdasarkan nilai pada **Tabel 5** sebagai berikut:

$$\begin{aligned}
APER &= \frac{n_{01} + n_{10}}{n_0 + n_1} \\
&= \frac{84 + 11}{101 + 299} \\
&= 0.2375
\end{aligned}$$

Hasil perhitungan APER menunjukkan bahwa model regresi probit biner yang diperoleh menghasilkan persentase kesalahan klasifikasi sebesar 0.2375 atau 23.75%.

$$\begin{aligned}
 \text{Hit Rasio} &= \frac{n_{00} + n_{11}}{n_0 + n_1} \times 100\% \\
 &= \frac{17 + 288}{101 + 299} \times 100\% \\
 &= 0.7625 \times 100\% \\
 &= 76.25\%
 \end{aligned}$$

Nilai *Hit Rasio* yang diperoleh mengartikan bahwa model regresi probit tiner yang terbentuk memiliki ketepatan klasifikasi sebesar 76.25%.

$$\begin{aligned}
 \text{Press's } Q &= \frac{[N - (rK)]^2}{N(K - 1)} \\
 &= \frac{[400 - (305 \times 2)]^2}{400(2 - 1)} \\
 &= 110.2500
 \end{aligned}$$

Hipotesis :

H_0 : pengklasifikasian tidak akurat

H_1 : pengklasifikasian akurat

Nilai *Press's Q* yang diperoleh sebesar 110.2500 sehingga nilai ini lebih besar dari nilai $\chi^2_{0.05,1} = 3.8415$ maka H_0 ditolak, hal ini memberikan kesimpulan bahwa pengklasifikasian akurat.

Hal ini menunjukkan bahwa pengklasifikasian status penerima bantuan PKH berdasarkan faktor jumlah anggota keluarga, status kepemilikan rumah tinggal, sumber air utama, dan kepemilikan televisi menggunakan regresi probit sudah tepat dengan tingkat keakuratan 76.25%

4. SIMPULAN

Hasil analisis dan pembahasan yang telah dilakukan, dapat ditarik kesimpulan sebagai berikut:

1. Model regresi probit pada status penerima bantuan sosial Program Keluarga Harapan adalah sebagai berikut:

$$\begin{aligned}
 Y^* &= -0.02428 + 0.63725x_1 + 0.41602x_2 - 0.02211x_3 + 0.11387x_4 + 0.09281x_5 \\
 &\quad + 0.15658x_6 + 0.27672x_7 + 0.59625x_8 + 0.07121x_9 - 0.08209x_{10} \\
 &\quad + 0.69366x_{11} + 1.02526x_{12}
 \end{aligned}$$

2. Faktor-faktor yang berpengaruh secara signifikan pada status penerima bantuan sosial Program Keluarga Harapan (PKH) adalah jumlah anggota keluarga (X_1), status kepemilikan rumah tinggal (X_2), sumber air utama (X_6), dan kepemilikan televisi (X_{12}).

3. Hasil ketepatan klasifikasi status penerima bantuan sosial Program Keluarga Harapan (PKH) adalah sebesar 76.25%. Hasil klasifikasi yang diperoleh sudah akurat berdasarkan nilai *Press's Q*
4. Hasil Penelitian ini diharapkan dapat memberikan masukan bagi pihak terkait agar masalah status penerima bantuan PKH dapat diminimalisir.

5. UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Universitas Mataram atas dukungan dan fasilitas yang telah diberikan, sehingga penelitian berjudul *Klasifikasi Status Penerima Bantuan PKH Menggunakan Regresi Probit* dapat terlaksana dengan baik.

6. REKOMENDASI

Penelitian selanjutnya diharapkan dapat mengeksplorasi perbandingan dengan pendekatan regresi biner alternatif, seperti model logit maupun decision tree regression, guna memperoleh model yang paling representatif dan optimal dalam merepresentasikan permasalahan yang dikaji.

7. REFERENSI

- Ariessela, S, G. R. dan P. I. (2021). Analisis Regresi Probit Biner Bivariat (Studi Kasus: Indeks Pendidikan dan Indeks Pengeluaran di Pulau Kalimantan Tahun 2017). *Jurnal EKSPONENSIAL*, 12(1), 73–82. <https://doi.org/https://doi.org/10.30872/ekspensial.v12i1.764>.
- Arimbi, Y, D. (2022). Kecemburuan Sosial Masyarakat Non Penerima Manfaat PKH di Dusun Pringroto EDsa Punjung Kecamatan Kebonagung Kabupaten Pacitan. *Jurnal Pengabdian Untuk Mu NegeRI*, 6(1), 163–167. <https://doi.org/https://doi.org/10.37859/jpumri.v6i1.3280>.
- Dewanti, C, R. V. dan R. A. T. (2019). Pemodelan Faktor-faktor yang Mempengaruhi Status Balita Stunting di Provinsi Jawa Timur Menggunakan Regresi Probit Biner. *Jurnal Sains Dan Seni ITS*, 8(2), 129–136. <https://doi.org/10.12962/j23373520.v8i2.48519>.
- Epriliyanti, Y. A dan Ratnasari, V. (2020). Pemodelan Faktor-faktor yang Mempengaruhi Keefektifan Sistem Pembelajaran Daring (SPADA) Menggunakan Regresi Probit Biner (Studi Kasus: Mahasiswa ITS Masa PAndemi Covid-19). *INFERENSI*, 3(2), 115–122. <https://doi.org/http://dx.doi.org/10.12962/j27213862.v3i2.7714>.
- Fathurahman, M. (2019). Pemodelan Indeks Pembangunan Kesehatan Masyarakat Kabupaten/Kota di Pulau Kalimantan Menggunakan Pendekatan Regresi Probit. *Jurnal Varian*, 2(2), 47–54. <https://doi.org/https://doi.org/10.30812/varian.v2i2.382>.
- Fatmallah, I, Baskara, Z. W, S. dan F. N. (2021). Analisis Faktor-faktor yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa MAtematika FMIPA Universitas Mataram Menggunakan Regresi Probit Biner. *Journal of Statistics*. <https://doi.org/http://eprints.unram.ac.id/id/eprint/43873>.
- Febyanti, F. (2022). Pemodelan Faktor-faktor yang Mempengaruhi Harga Rumah di Jabodetabek Menggunakan Metode Regresi Probit. *Jurnal Riset Statistika*, 2(1), 51–58. <https://doi.org/https://doi.org/10.29313/jrs.vi.905>.
- Gentara, R, R. R. dan S. (2023). Efektivitas Program Keluarga Harapan (PKH) dalam

- Meningkatkan Sosial Ekonomi Masyarakat Desa. *Jurnal Ekonomi Dan Bisnis*, 11(3), 373–385. <https://doi.org/https://doi.org/10.58406/jeb.v11i3.1370>.
- Harsyiah, L, Hadijati, M, dan F. N. (2024). Perbandingan Analisis Diskriminan dan Naive Bayes dalam Pengklasifikasian Status Penerima Bantuan Program Keluarga Harapan di NTB. *Jurnal Matematika UNAND*, 13(4), 296–308. <https://doi.org/https://doi.org/10.25077/jmua.13.4.296-308.2024>
- Johnson, R. A. dan Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall, Inc.
- Manik, A. (2020). Penerapan MOORA dalam Pendukung Keputusan Kelayakan Penerimaan Bantuan Program Keluarga Harapan (PKH). *Jurnal Sistem Komputer Dan Informatika*, 2(1), 42–47. <https://doi.org/http://dx.doi.org/10.30865/json.v2i1.2469>
- NTB, S. (2024). Selama 2023 Capai Rp. 945,6 Miliar, Bansos PKH Jadi Salah Satu Penurunan Angka Kemiskinan di NTB. *Suarantb.Com*. <https://www.suarantb.com/2024/01/30/selama-2023-capai-rp9456-miliar-bansos-pkh-jadi-salah-satu-penurun-angka-kemiskinan-di-ntb/>.
- Nur'eni dan Handayani, L. (2020). Regresi Probit untuk Analisis Variabel-variabel yang Mempengaruhi Perceraian di Sulawesi Tengah. *Jurnal Aplikasi Statistika Dan Komputasi*, 12(1), 13–21. <https://doi.org/https://doi.org/10.34123/jurnalasks.v12i1.211>
- Romli, I dan Zy, T, A. (2020). Penentuan Jadwal Overtime dengan Klasifikasi Data Karyawan Menggunakan Algoritma C4.5. *Jurnal Sains Komputer Dan Informatika*, 4(2), 694–702. <https://doi.org/http://dx.doi.org/10.30645/j-sakti.v4i2.260>
- Sianipar, P. N, dan Cipta, H. (2023). Sistem Pendukung Keputusan Penerima Bantuan Sosial Program Keluarga Harapan (PKH) Kelurahan Titi Kuning dengan Metode VIKOR. *Jurnal Riset Sistem Informasi Dan Teknik Informatika*, 8(1), 18–27. <https://doi.org/http://dx.doi.org/10.30645/jurasik.v8i1.537.g515>