

MODEL REGRESI LOGISTIK ROBUST PADA STATUS HIPERKOLESTEROLEMIA MENGGUNAKAN ESTIMASI BIANCO-YOHAI

Retno Wulandari, Sri Sulistijowati Handajani, dan Diari Indriati
Program Studi Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Sebelas Maret

ABSTRAK. Hiperkolesterolemia merupakan gangguan metabolisme kolesterol yang disebabkan oleh kadar kolesterol dalam darah yang berlebih. Hiperkolesterolemia dapat dipengaruhi oleh kadar *HDL*, *LDL*, dan trigliserida. Penelitian ini dilakukan untuk mengetahui hubungan antara ketiga faktor dengan status hiperkolesterolemia. Status hiperkolesterolemia berupa kategori yaitu positif dan negatif sehingga digunakan model regresi logistik. Akan tetapi, karena terdapat pencilan pada data maka digunakan model regresi logistik *robust* dengan metode estimasi Bianco-Yohai. Berdasarkan hasil penelitian didapatkan 4 pencilan terhadap variabel respon Y dan 6 pencilan terhadap variabel prediktor X . Hasil estimasi Bianco-Yohai diperoleh model regresi logistik dengan $\hat{\beta}$ masing-masing yaitu $\hat{\beta}_0 = -10.2140$, $\hat{\beta}_1 = -0.0265$, $\hat{\beta}_2 = 0.0673$, dan $\hat{\beta}_3 = 0.0107$.

Kata Kunci: model regresi logistik, pencilan, Bianco-Yohai, hiperkolesterolemia

1. PENDAHULUAN

Kolesterol adalah zat lemak atau lipid yang fungsinya sangat penting bagi tubuh. Kolesterol dapat diproduksi oleh tubuh ataupun dihasilkan dari sumber makanan. Pada kadar tertentu kolesterol mempunyai banyak manfaat namun apabila kadar kolesterol terlalu tinggi dapat menyebabkan hiperkolesterolemia. Hiperkolesterolemia merupakan gangguan metabolisme kolesterol yang disebabkan oleh kadar kolesterol dalam darah yang berlebih (Guyton dan Hall [6]).

Individu dikatakan positif hiperkolesterolemia apabila kadar total kolesterol dalam darah berada di atas batas normal yaitu > 200 mg/dL. Sedangkan individu dikatakan negatif hiperkolesterolemia apabila kadar total kolesterol dalam darah berada pada batas normal yaitu ≤ 200 mg/dL (Fatimah [5]). Hiperkolesterolemia dapat dipengaruhi oleh berapa faktor sehingga diperlukan suatu model untuk menentukan hubungannya yaitu model regresi. Akan tetapi pada kasus ini variabel Y berupa kualitatif dengan dua kategori yaitu positif hiperkolesterolemia dan negatif hiperkolesterolemia sehingga digunakan model regresi logistik (Hosmer dan Lemeshow [7]).

Pada saat membentuk model regresi logistik dapat dilakukan deteksi terhadap adanya pencilan pada data. Pencilan adalah pengamatan yang jauh dari pusat data yang mungkin berpengaruh besar terhadap koefisien regresi (Sembiring [11]). Keberadaan pencilan tidak dapat diabaikan dikarenakan apabila terdapat pencilan dapat memengaruhi kesimpulan akhir dari model yang terpilih.

Pada tahun 1996, Bianco dan Yohai [2] memperkenalkan metode yang digunakan untuk mengestimasi parameter regresi logistik agar menghasilkan model regresi logistik *robust* yang baik digunakan apabila ditemukan pencilan pada data. Kemudian beberapa penelitian yang dilakukan setelahnya menyebut metode tersebut dengan metode estimasi Bianco-Yohai.

Penelitian tentang estimasi parameter model regresi untuk penentuan status hiperkolesterolemia dilakukan oleh Fatimah [5] dengan metode estimasi *MLE* diperoleh yaitu status hiperkolesterolemia dipengaruhi oleh tiga faktor yaitu kadar *low density lipoprotein (LDL)*, *high density lipoprotein (HDL)*, dan trigliserida. Penelitian lain yang masih berhubungan dengan hiperkolesterolemia dilakukan oleh Lirani [9]. Akan tetapi pada penelitian tersebut tidak dibentuk model regresi dan tidak dilakukan deteksi adanya pencilan pada data. Selanjutnya dengan menggunakan data sekunder dari hasil penelitian Lirani [9], penelitian akan dikembangkan dengan membentuk model regresi logistik *robust* menggunakan estimasi Bianco-Yohai dikarenakan data pada penelitian tersebut mengandung pencilan.

2. MODEL REGRESI LOGISTIK

Model regresi logistik adalah model regresi dengan variabel respon hanya memiliki dua kemungkinan hasil misalnya sukses dan tidak sukses (Hosmer dan Lemeshow [7]. Variabel respon tersebut biasanya diberi kode 1 (untuk menyatakan kejadian sukses atau *event*) dan 0 (untuk menyatakan kejadian tidak sukses atau *non-event*).

Menurut Agresti [1] model logistik dengan m variabel prediktor dinyatakan sebagai

$$P(x) = \left(\frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im} + \varepsilon_i)}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im} + \varepsilon_i)} \right), \quad (2.1)$$

dengan β merupakan parameter regresi, x merupakan variabel prediktor, dan $i = 1, \dots, m$.

Model logistik pada Persamaan 2.1 ditransformasikan dengan menggunakan transformasi logit sehingga diperoleh

$$\begin{aligned} \log \left(\frac{P(x)}{1 - P(x)} \right) &= \log(\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_m x_{im} + \varepsilon_i)) \\ &= \beta_0 + \beta_1 x_{i1} + \dots + \beta_m x_{im} + \varepsilon_i. \end{aligned}$$

3. PENDETEKSIAN PENCILAN

Pada penelitian ini digunakan dua metode untuk mendeteksi adanya pencilan yaitu *hat matrix* dan analisis *deviance residual*.

3.1. **Hat Matrix.** Metode *hat matrix* digunakan untuk mendeteksi pencilan pada variabel prediktor X yang didefinisikan sebagai $\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$. Metode pendeteksiannya yaitu dengan menghitung nilai h_{ii} yang merupakan unsur ke- i pada diagonal utama *hat matrix* yaitu

$$h_{ii} = \mathbf{x}_i (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_i^T,$$

dengan $\mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{m1} \\ 1 & x_{12} & x_{22} & \cdots & x_{m2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{mn} \end{pmatrix}$ dan \mathbf{x}_i adalah vektor baris dari penga-

matan ke- i . Suatu pengamatan ke- i merupakan pencilan apabila $h_{ii} > \frac{2m}{n}$ dengan m merupakan banyaknya parameter dan n merupakan banyaknya pengamatan (Rousseeuw dan Leroy [10]).

3.2. **Analisis Deviance Residual.** Metode analisis *deviance residual* digunakan untuk mendeteksi pencilan pada respon Y . Menurut Agresti [1], *residual* diartikan sebagai ukuran kesesuaian antara suatu pengamatan pada variabel respon dan nilai estimasinya. *Residual* pada pengamatan ke- i diperoleh menggunakan rumus $y_i - \hat{P}_i$ dengan y_i adalah nilai respon 0 atau 1 dan \hat{P}_i adalah nilai estimasinya. *Deviance residual* pengamatan ke- i dari vektor respon dapat dinyatakan oleh Persamaan 3.1

$$d_i = \begin{cases} \left[2y_i \log \left(\frac{y_i}{\hat{P}_i} \right) + 2(n - y_i) \log \left(\frac{(n - y_i)}{n - \hat{P}_i} \right) \right]^{1/2}, & y_i \geq \hat{P}_i \\ - \left[2y_i \log \left(\frac{y_i}{\hat{P}_i} \right) + 2(n - y_i) \log \left(\frac{(n - y_i)}{n - \hat{P}_i} \right) \right]^{1/2}, & y_i < \hat{P}_i \end{cases} \quad (3.1)$$

dengan n adalah banyaknya pengamatan. Suatu pengamatan ke- i merupakan pencilan apabila $|d_i| > 2$ (Jaelani [8]).

4. METODE *Maximum Likelihood Estimation (MLE)*

Untuk mengestimasi parameter β pada regresi logistik, metode estimasi yang digunakan adalah *maximum likelihood estimation (MLE)* dengan menggunakan fungsi *likelihood* yaitu

$$L = \prod_{i=1}^n (P_i)^{y_i} + (1 - P_i)^{1 - y_i}.$$

dengan P_i merupakan probabilitas respon ke- i . Metode *MLE* memberikan estimasi dari β dengan memaksimalkan fungsi log-*likelihood*-nya (Hosmer dan Lemeshow [7]). Fungsi log-*likelihood* yang dimaksud yaitu

$$l = \sum_{i=1}^n \left[y_i \log(P_i) + (1 - y_i) \log(1 - P_i) \right],$$

yang dapat digunakan untuk mengestimasi β . Selain itu, dengan mempertimbangkan *deviance* yang diperoleh dari fungsi *log-likelihood* regresi logistik yaitu

$$D_i(\beta) = -[y_i(\log(\pi_i)) + (1 - y_i)(\log(1 - \pi_i))], \quad (4.1)$$

sehingga *MLE* juga dapat didefinisikan dengan meminimumkan $\sum_{i=1}^n D_i(\beta)$ (Casiawaty [3]).

5. ESTIMASI BIANCO-YOHAI

Bianco dan Yohai [2] menyatakan bahwa *MLE* sangat sensitif terhadap data pencilan dalam suatu sampel pengamatan. *MLE* tidak *robust* dalam menghadapi pencilan khususnya pada regresi logistik. *MLE* didefinisikan dengan meminimumkan persamaan 4.1

Bianco dan Yohai [2] mendefinisikan metode estimasi yang lebih *robust* digunakan untuk mengestimasi parameter model regresi logistik pada data yang mengandung pencilan dengan meminimumkan

$$l_{BY}(\beta) = \sum_{i=1}^n [\rho(D_i(\beta)) + (G(P_i)) + (G(1 - P_i))]. \quad (5.1)$$

$G(t)$ adalah suatu fungsi yang didefinisikan sebagai

$$G(t) = \int_0^t \psi(-\ln u) du,$$

dengan $\psi(t) = \rho'(t)$, $\psi(t) \geq 0$. Fungsi $\rho(t)$ merupakan fungsi terbatas, dapat diturunkan, dan merupakan fungsi tidak turun yang didefinisikan sebagai

$$\rho(t) = \begin{cases} t - \frac{t^2}{k}, & \text{jika } t \leq k \\ \frac{k}{2}, & \text{jika } t > k \end{cases}$$

Nilai $k > -\ln(\frac{1}{2})$ sedemikian sehingga $\psi(t) > 0$ untuk setiap $0 < t < k$ diperoleh

$$G(t) = \begin{cases} 0, & \text{jika } t \leq e^{-k} \\ \int_0^t t + (\frac{t \ln t}{k}), & \text{jika } t > e^{-k} \end{cases}$$

6. UJI SIGNIFIKANSI PARAMETER

Uji signifikansi parameter dilakukan secara serempak dan parsial. Uji signifikansi secara serempak dilakukan menggunakan uji G dengan hipotesis yaitu $H_0 : \beta_1 = \beta_2 = \dots = \beta_m = 0$ (tidak ada pengaruh variabel prediktor terhadap variabel respon) dan $H_1 : \exists \beta_j \neq 0$, dengan $j = 1, 2, \dots, m$ (paling tidak terdapat satu variabel prediktor yang berpengaruh terhadap variabel respon). Statistik uji yang digunakan adalah $G = -2(L_0 - L_1)$ dengan L_0 merupakan *log-likelihood* tanpa prediktor dan L_1 merupakan *log-likelihood* dengan prediktor.

Uji signifikansi secara parsial dilakukan menggunakan uji Wald dengan hipotesis yaitu $H_0 : \beta_j = 0$ untuk suatu j , dengan $j = 1, 2, \dots, m$ (tidak ada pengaruh variabel prediktor ke- j terhadap variabel respon), dan $H_1 : \beta_j \neq 0$ untuk suatu j , dengan $j = 1, 2, \dots, m$ (ada pengaruh variabel prediktor ke- j terhadap variabel respon). Statistik uji yang digunakan adalah

$$W = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right)^2,$$

dengan $\hat{\beta}_j$ merupakan estimasi dari β_j dan $SE(\hat{\beta}_j)$ merupakan standar error dari $\hat{\beta}_j$. Uji Wald menggunakan daerah kritis yaitu $DK = \{W | W > \chi^2_{(\alpha, m)}\}$. H_0 ditolak jika $W \in DK$ pada taraf signifikansi α (Hosmer dan Lemeshow [7]).

7. UJI KELAYAKAN MODEL

Uji kelayakan model dilakukan untuk mengetahui apakah model layak digunakan atau tidak. Uji yang digunakan adalah uji Hosmer dan Lemeshow dengan hipotesis yaitu H_0 : model layak digunakan dan H_1 : model tidak layak digunakan. Statistik uji yang digunakan adalah

$$\hat{C} = \sum_{z=1}^g \frac{(o_z - n'_z \bar{\pi}_z)^2}{n'_z \bar{\pi}_z (1 - \bar{\pi}_z)},$$

dengan n'_z merupakan total pengamatan pada kelompok ke- z , o_z merupakan banyaknya pengamatan kejadian sukses dalam kelompok ke- z , dan $\bar{\pi}_z$ merupakan rata-rata estimasi probabilitas sukses kelompok ke- z . $\sum_{z=1}^g n'_z = n$, dengan $z = 1, 2, \dots, g$. Uji Hosmer dan Lemeshow menggunakan daerah kritis yaitu $DK = \{\hat{C} | \hat{C} > \chi^2_{(\alpha, g-2)}\}$. H_0 ditolak jika $\hat{C} \in DK$ pada taraf signifikansi α (Hosmer dan Lemeshow [7]).

8. HASIL DAN PEMBAHASAN

Pada artikel ini dibahas model regresi logistik, pendeteksian pencilan, dan model regresi logistik *robust* beserta ujinya.

8.1. Pendeteksian Pencilan. Pendeteksian pencilan terhadap variabel respon Y dilakukan menggunakan analisis *deviance residual* yaitu dengan menghitung nilai *deviance residual* (d_i). Suatu pengamatan merupakan pencilan apabila nilai $|d_i| > 2$. Hasil yang diperoleh adalah data ke-27, ke-28, ke-29, dan ke-30 merupakan pencilan terhadap variabel respon Y . Pendeteksian pencilan terhadap X dilakukan menggunakan *hat matrix* yaitu dengan menghitung nilai h_{ii} . Jika $h_{ii} > \frac{2m}{n}$ yaitu $h_{ii} > 0.075$ maka merupakan pencilan terhadap variabel prediktor

X . Hasil yang diperoleh adalah data ke-9, ke-11, ke-19, ke-35, ke-61, dan ke-76 merupakan pencilan terhadap variabel prediktor X .

8.2. Model Regresi Logistik *Robust*. Estimasi regresi logistik *robust* yang digunakan dalam penelitian ini adalah estimasi Bianco-Yohai. Nilai awal yang digunakan yaitu hasil dari estimasi menggunakan *MLE*, disajikan pada Tabel 1

Tabel 1. Hasil estimasi parameter menggunakan *MLE*.

Parameter	Koefisien	Standar Error
β_0	-10.9275	3.0546
β_1	-0.0095	0.0392
β_2	0.0684	0.01843
β_3	0.0103	0.0068

Hasil estimasi parameter menggunakan metode Bianco-Yohai berdasarkan nilai awal pada Tabel 1 dapat dilihat pada Tabel 2.

Tabel 2. Hasil estimasi parameter menggunakan metode Bianco-Yohai.

Parameter	Koefisien	Standar Error
β_0	-10.2140	2.9029
β_1	-0.0265	0.0269
β_2	0.0673	0.0186
β_3	0.0107	0.0049

Berdasarkan Tabel 1 dan Tabel 2 dapat disimpulkan bahwa estimasi menggunakan metode Bianco-Yohai menghasilkan parameter dengan standar error yang lebih kecil sehingga lebih baik untuk digunakan.

Berdasarkan parameter yang telah diperoleh menggunakan estimasi Bianco-Yohai dapat dibentuk model regresi logistik *robust* yaitu

$$P(x) = \left(\frac{\exp(-10.2140 - 0.0265x_{i1} + 0.0673x_{i2} + 0.0107x_{i3} + \varepsilon_i)}{1 + \exp(-10.2140 - 0.0265x_{i1} + 0.0673x_{i2} + 0.0107x_{i3} + \varepsilon_i)} \right), \quad (8.1)$$

dengan $P(x)$ adalah variabel respon Y yang berupa probabilitas untuk pengamatan ke- i , x_{1i} adalah variabel prediktor X_1 pengamatan ke- i , x_{2i} adalah variabel prediktor X_2 pengamatan ke- i , x_{3i} adalah variabel prediktor X_3 pengamatan ke- i , dan ε_i adalah residu ke- i dimana $i = 1, 2, 3, \dots, 80$.

8.3. Uji Signifikansi Parameter. Parameter yang telah diperoleh menggunakan estimasi Bianco-Yohai selanjutnya diuji kesignifikansiannya. Uji signifikansi dilakukan secara serempak dan secara parsial menggunakan uji G dan uji Wald.

- (1) Uji Serempak. Hipotesis $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ (tidak ada pengaruh variabel prediktor terhadap variabel respon). Hipotesis $H_1 : \exists \beta_j \neq 0$, dengan $j = 1, 2, 3$ (paling tidak terdapat satu variabel prediktor yang berpengaruh

terhadap variabel respon). Menggunakan taraf signifikansi $\alpha = 0.05$ dan daerah kritis $DK = \{G|G \geq \chi_{(0.05,3)}^2\} = \{G|G \geq 7.81472\}$, H_0 ditolak jika $G \in DK$. Statistik uji pada pengujian ini yaitu $G = 45.748$. Berdasarkan statistik uji yaitu $G \in DK$ sehingga diperoleh kesimpulan yaitu H_0 ditolak yang berarti bahwa paling tidak terdapat satu variabel prediktor yang berpengaruh terhadap variabel respon. Selanjutnya dilakukan Uji signifikansi masing-masing variabel.

- (2) Uji Parsial. Hipotesis $H_0 : \beta_j = 0$, dengan $j = 1, 2, 3$ (tidak ada pengaruh variabel prediktor ke- j terhadap variabel respon). Hipotesis $H_1 : \beta_j \neq 0$, dengan $j = 1, 2, 3$ (ada pengaruh variabel prediktor ke- j terhadap variabel respon). Menggunakan taraf signifikansi $\alpha = 0.05$ dan daerah kritis $DK = \{W|W > \chi_{(0.05,3)}^2\} = \{W|W > 7.81472\}$, H_0 ditolak jika $W \in DK$. Statistik uji pada pengujian ini yaitu W pada Tabel 3. Berdasarkan statis-

Tabel 3. Nilai W masing-masing variabel.

Variabel	W	$\chi_{(0.05,3)}^2$	Kesimpulan
X_1	0.966118469	7.81472	H_0 tidak ditolak
X_2	13.09270455	7.81472	H_0 ditolak
X_3	4.729650521	7.81472	H_0 tidak ditolak

tik uji, diperoleh kesimpulan yaitu H_0 ditolak untuk variabel X_2 sehingga X_2 (LDL) berpengaruh signifikan terhadap variabel respon (Y) atau status hiperkolesterolemia dan H_0 tidak ditolak untuk variabel X_1 dan X_3 sehingga variabel X_1 (HDL) dan X_3 (trigliserida) tidak berpengaruh signifikan terhadap variabel respon (Y) atau status hiperkolesterolemia.

8.4. Uji Hosmer dan Lemeshow. Setelah dilakukan uji signifikansi masing-masing parameter kemudian dilakukan uji kelayakan model menggunakan uji Hosmer dan Lemeshow. Uji ini dilakukan untuk mengetahui apakah model layak digunakan pada data atau tidak. Hipotesis H_0 : Model layak digunakan. Hipotesis H_1 : Model tidak layak digunakan. Menggunakan taraf signifikansi $\alpha = 0.05$ dan daerah kritis $DK = \{\hat{C}|\hat{C} > \chi_{(0.05,8)}^2\} = \{\hat{C}|\hat{C} > 15.5071\}$, H_0 ditolak jika $\hat{C} \in DK$. Statistik uji pada pengujian ini yaitu \hat{C} pada Tabel 4. Berdasarkan

Tabel 4. Nilai \hat{C} .		
\hat{C}	$\chi_{(0.05,8)}^2$	Kesimpulan
15.1690	15.5071	H_0 tidak ditolak

statistik uji, diperoleh kesimpulan yaitu H_0 tidak ditolak sehingga model layak digunakan.

9. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dilakukan, diperoleh kesimpulan yaitu model regresi logistik *robust* pada status hiperkolesterolemia disajikan sebagai

$$P(x) = \left(\frac{\exp(-10.2140 - 0.0265x_{i1} + 0.0673x_{i2} + 0.0107x_{i3} + \varepsilon_i)}{1 + \exp(-10.2140 - 0.0265x_{i1} + 0.0673x_{i2} + 0.0107x_{i3} + \varepsilon_i)} \right).$$

Nilai odds ratio untuk kadar *HDL* yaitu 0.00000622 artinya setiap kenaikan 1 persen kadar *HDL* memiliki resiko positif hiperkolesterolemia sebesar 0.00000622 kali. Nilai odds ratio untuk kadar *LDL* yaitu 0.99953814 artinya setiap kenaikan 1 persen kadar *LDL* memiliki resiko positif hiperkolesterolemia sebesar 0.99953814 kali. Nilai odds ratio untuk kadar trigliserida yaitu 0.00493622 artinya setiap kenaikan 1 persen kadar trigliserida memiliki resiko positif hiperkolesterolemia sebesar 0.00493622 kali. Berdasarkan nilai odd ratio ketiga variabel dapat disimpulkan bahwa semakin bertambah kadar *HDL*, *LDL*, dan trigliserida dalam darah akan meningkatkan resiko hiperkolesterolemia.

DAFTAR PUSTAKA

- [1] Agresti, A., *Categorical Data Analysis*, 2ed., John Wiley and Sons, New York, 2002.
- [2] Bianco, A.M. and V.J. Yohai, *Robust Estimation in the Logistic Regression Model*, Robust Statistics, Data Analysis, and Computer Intensive Methods (New York) (H. Rieder ed.), Springer, 1996, 17-34.
- [3] Casiawaty, L., *Pemodelan Regresi Logistik Robust pada Data Outlier Menggunakan Penduga Bianco-Yohai*, UNHAS Repository, 2015, 1-10.
- [4] Croux, C. and G. Haesbroeck, *Implementing the Bianco and Yohai Estimator for Logistic Regression*, Computational Statistics and Data Analysis **44** (2003), 273-295.
- [5] Fatimah, S.N., *Estimasi Parameter Model Regresi Logistik Menggunakan Metode Residual Bootstrap*, Skripsi, Universitas Gadjah Mada, Yogyakarta, 2000.
- [6] Guyton, A.C., and Hall, J.E. *Buku Ajar Fisiologi Kedokteran*, 11 ed., EGC, Jakarta, 2008
- [7] Hosmer, D.W. and S. Lemeshow, *Applied Logistic Regression*, 2 ed., John Wiley and Sons, New York, 2000.
- [8] Jaelani, U., *Pendeteksian Outlier dalam Regresi Logistik Biner dengan Metode Detects Outliers Using Weight (DOUW)*, Tesis, Universitas Padjajaran, Bandung, 2010.
- [9] Lirani, E.A., *Hubungan Kontrol Glikemik (HbA1c) dengan Kejadian Nefropatik Diabetik Pasien Diabetes melitus Tipe 2 Di RSUD Dr. Moewardi Surakarta*, Skripsi. Universitas Sebelas Maret, Surakarta, 2015.
- [10] Rousseeuw, P.J. dan A.M. Leroy, *Robust Regression and Outlier Detection*, John Wiley and Sons, New York, 1987, 216-247.
- [11] Sembiring, R.K., *Analisis Regresi*, 2 ed., Institut Teknologi Bandung, Bandung, 2003.
- [12] Soemartini, *Pencilan (Outlier)*, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Padjajaran, Bandung, 2007.