

Penerapan *Maximum Likelihood Estimator* dalam Model Regresi Logistik pada Data Penderita Penyakit Akibat Infeksi Virus *Dengue*

Safira Salsabila*, Dwi Agustin Nuriani Sirodj

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

*safirasalsabila46@gmail.com, dwi.agustin@gmail.com

Abstract. Linear regression analysis is used to determine the relationship of predictor variables and response variable. However, there are response variables in categorical or binary so linear regression analysis method cannot be used because it breaks the assumptions. To solve this problem, logistic regression analysis is used using the maximum likelihood estimator (MLE). Dengue Fever (DD) and Dengue Hemorrhagic Fever (DHF) are diseases caused by the dengue virus which is transmitted by the bite of *Aedes aegypti* dan *Aedes albopictus*. Although caused by the same virus, but the consequences or effects caused have different severity. In this study, we want to know the logistic regression model with MLE method by using reseach data by Widya N.M. Sari in 2017 which is data of medical record for patients with Dengue Fever and Dengue Hemorrhagic Fever in December 2015 – February 2016 at Haji Hospital Surabaya. The data consists of response variable is classification of patients and predictor variables are the difference level of leukocytes, platelets, hematocrit, and hemoglobin from the first and third day of medical check up. The result is that binary logistic regression model is feasible to use with the parameter estimation are $\beta_0 = -1,569552$, $\beta_1 = -0,3058617$, and $\beta_2 = 0,1862561$. The predictor variables that significant are the difference level of leukocytes and platelets from the first and third day of medical check up.

Keywords: Logistic Regression, *Maximum Likelihood Estimator* (MLE), Dengue Hemorrhagic Fever (DHF).

Abstrak. Analisis regresi linear digunakan untuk mengetahui hubungan pengaruh variabel prediktor terhadap variabel respon. Akan tetapi, banyak data dengan variabel responnya berbentuk kategorik atau biner sehingga metode analisis regresi linear tidak dapat digunakan karena melanggar asumsi. Untuk mengatasi permasalahan tersebut digunakan analisis regresi logistik menggunakan *maximum likelihood estimator* (MLE). Penyakit Demam *Dengue* (DD) dan Demam Berdarah *Dengue* (DBD) adalah penyakit yang disebabkan oleh virus *dengue* yang ditularkan oleh gigitan nyamuk *Aedes aegypti* dan *Aedes albopictus*. Walaupun disebabkan oleh virus yang sama, akan tetapi akibat yang ditimbulkan memiliki tingkat keparahan yang berbeda. Pada penelitian ini ingin mengetahui model regresi logistik dengan metode MLE dengan menggunakan data penelitian Widya N.M. Sari tahun 2017 yang merupakan data rekam medis pada penderita DD dan DBD bulan Desember 2015 – Februari 2016 dari Rumah Sakit Umum Haji Surabaya. Data tersebut terdiri dari variabel responnya adalah klasifikasi penyakit penderita dan variabel prediktornya adalah selisih kadar leukosit, trombosit, hematokrit, dan hemoglobin dari pemeriksaan hari pertama dan ketiga. Hasil penelitian didapat model regresi logistik biner yang layak digunakan dengan estimasi parameternya adalah $\beta_0 = -1,569552$, $\beta_1 = -0,3058617$, dan $\beta_2 = 0,1862561$. Variabel prediktor yang berpengaruh adalah selisih kadar leukosit dan trombosit pemeriksaan hari pertama dan ketiga.

Kata Kunci: Regresi Logistik, *Maximum Likelihood Estimator* (MLE), Demam

Berdarah *Dengue* (DBD).

1. Pendahuluan

Analisis regresi linear telah digunakan pada berbagai bidang. Analisis regresi linear digunakan untuk mengetahui hubungan pengaruh variabel prediktor terhadap variabel respon. Untuk mengestimasi parameter, regresi linear menggunakan metode kuadrat terkecil biasa atau *ordinary least square* (OLS). Pada regresi linear, variabel prediktor dan variabel respon berbentuk numerik. Akan tetapi, di lapangan terdapat data dengan variabel respon berbentuk kategorik atau biner. Sehingga tidak dapat menggunakan analisis regresi linear karena melanggar asumsi, yaitu terdapat masalah non-linearitas dan heteroskedastisitas. Oleh karena itu, untuk mengatasi permasalahan tersebut dengan menggunakan pendekatan kemungkinan maksimum (*maximum likelihood*) berdasarkan pada fungsi kemungkinan dari distribusi bernoulli atau binomial sebagai dasar pada pemodelan regresi logistik (Hajarisman, 2009).

Penyakit Demam *Dengue* (DD) dan Demam Berdarah *Dengue* (DBD) disebabkan oleh virus *dengue* yang ditularkan oleh gigitan nyamuk *Aedes aegypti* dan *Aedes albopictus* (Sari, 2017). Walaupun disebabkan oleh virus yang sama dan ditandai dengan adanya demam, tetapi dua penyakit ini memiliki perbedaan. Perbedaan dua penyakit ini yaitu pada tingkat keparahan dari penyakit (Kemenkes, 2017). Sebagai penyakit yang memiliki resiko kematian, maka dilakukan pemeriksaan terhadap pasien untuk mendeteksi apakah pasien akan mengalami Demam Berdarah *Dengue* (DBD) atau tidak.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: “Bagaimana model regresi logistik biner dengan menggunakan *maximum likelihood estimator* (MLE) pada data penderita penyakit akibat infeksi virus *dengue*?”. Selanjutnya, tujuan dalam penelitian ini adalah untuk memodelkan regresi logistik biner dengan menggunakan *maximum likelihood estimator* (MLE) pada data penderita penyakit akibat infeksi virus *dengue*. Penerapannya menggunakan data rekam medis pada penderita Demam *Dengue* (DD) dan penderita Demam Berdarah *Dengue* (DBD) bulan Desember 2015 – Februari 2016. Data ini merupakan data sekunder dari penelitian yang dilakukan oleh Widya N.M. Sari tahun 2017 yang mana datanya diperoleh dari Rumah Sakit Umum (RSU) Haji Surabaya.

2. Metodologi

Data dan Sumber Data

Data yang digunakan dalam penelitian adalah data sekunder dari penelitian yang dilakukan oleh Widya N.M. Sari tahun 2017 yang mana datanya merupakan data rekam medis pada penderita penyakit Demam *Dengue* (DD) dan penderita penyakit Demam Berdarah *Dengue* (DBD) bulan Desember 2015 – Februari 2016 di Rumah Sakit Umum Haji Surabaya. Jumlah pengamatan pada data sebanyak 70 pasien. Data terdiri dari variabel respon berupa klasifikasi penyakit penderita dan variabel prediktor berupa selisih kadar leukosit, trombosit, hematokrit, dan hemoglobin pada pemeriksaan hari pertama dan hari ketiga.

Tabel 1. Variabel Penelitian

Variabel	Nama Variabel	Keterangan (Satuan)	Skala
X ₁	Selisih Kadar Leukosit	10 ³ sel/mm ³	Rasio
X ₂	Selisih Kadar Trombosit	10 ⁴ keping/mm ³	Rasio
X ₃	Selisih Kadar Hematokrit	%	Rasio
X ₄	Selisih Kadar Hemoglobin	g/dL	Rasio
Y	Klasifikasi Penyakit Penderita	0 : Penderita DD 1 : Penderita DBD	Nominal

Prosedur Penelitian

Langkah-langkah yang dilakukan dalam penelitian ini, yaitu:

1. Syarat memulai analisis regresi logistik biner adalah memiliki data dengan variabel responnya berbentuk kategorik atau biner. Selanjutnya dapat menentukan model regresi logistik binernya. Untuk menentukan model regresi logistik biner, nilai parameter β ditaksir menggunakan metode MLE. Dalam Sari (2017) menjelaskan variabel respon Y_i berdistribusi Bernoulli dengan *Probability Density Function* (PDF) sebagai berikut:

$$f(y_i) = \pi_i^{y_i}(1 - \pi_i)^{1-y_i}; y_i = 0, 1 \quad i = 1, 2, \dots, n$$

Kemudian dicari fungsi *likelihood*nya sebagai berikut:

$$L(\beta) = \prod_{i=1}^n f(y_i) = \prod_{i=1}^n \pi_i^{y_i}(1 - \pi_i)^{1-y_i}$$

Kemudian dicari fungsi *log-likelihood*nya sebagai berikut:

$$l(\beta) = \ln L(\beta) = \sum_{i=1}^n [y_i \ln \pi_i + (1 - y_i) \ln (1 - \pi_i)]$$

Kemudian, persamaan di atas dimaksimumkan untuk perhitungan estimasi parameter β . Syarat agar fungsi *log-likelihood* mencapai maksimum adalah $\partial l(\beta)/\partial \beta = 0$. Karena turunan tersebut merupakan fungsi nonlinear dalam β , maka untuk mengestimasi menggunakan metode Newton-Raphson. Dalam Nugraha (2013) menjelaskan model regresi logistik binernya sebagai berikut:

$$\pi_i = \left(\frac{\exp(X_i \beta)}{1 + \exp(X_i \beta)} \right)$$

π_i = probabilitas terjadinya DBD pada pengamatan ke- i

$X_i = (1, x_{i1}, \dots, x_{i4})$ = vektor dari nilai-nilai variabel prediktor pengamatan ke- i

$\beta = (\beta_0^*, \beta_1^*, \dots, \beta_4^*)'$ = vektor parameter pada regresi logistik

2. Setelah mendapatkan model, dilakukan pengujian parameter secara simultan dan secara parsial, dan pengujian kelayakan model.

- a. Uji parameter secara simultan dilakukan untuk mengetahui penduga parameter secara bersamaan apakah mempengaruhi variabel respon. Pengujian ini menggunakan statistik uji G. Apriliyanti (2021) menjelaskan pengujiannya sebagai berikut:

Hipotesis

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

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

Statistik Uji

$$G = -2(L_0 - L_k)$$

L_0 = nilai log likelihood model regresi logistik tanpa variabel prediktor

L_k = nilai log likelihood model regresi logistik dengan variabel prediktor

Kriteria Uji

Tolak H_0 jika nilai $G > \chi^2_{(\alpha; k)}$ dengan menggunakan $\alpha = 5\%$ dan k adalah banyaknya variabel prediktor

- b. Uji parameter secara parsial dilakukan untuk mengetahui penduga parameter secara individu apakah mempengaruhi variabel respon. Pengujian menggunakan statistik uji *Wald*. Kotimah & Wulandari (2014) menjelaskan pengujiannya sebagai berikut:

Hipotesis

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0$$

$$j = 1, 2, \dots, k$$

Statistik Uji

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

$\hat{\beta}_j$ = estimasi dari β_j
 $SE(\hat{\beta}_j)$ = standard error bagi β_j

Kriteria Uji

Tolak H_0 jika nilai $|W| > Z_{(\alpha/2)}$ dengan menggunakan $\alpha = 5\%$

- c. Uji kelayakan model dilakukan untuk menguji model yang didapat apakah layak digunakan. Pengujian ini menggunakan statistik uji Hosmer Lemeshow. Wulandari (2018) menjelaskan pengujiannya sebagai berikut:

Hipotesis

H_0 : Model layak digunakan

H_1 : Model tidak layak digunakan

Statistik Uji

$$\hat{C} = \sum_{z=1}^g \frac{(O_z - n'_z \pi_z)^2}{n'_z \pi_z (1 - \pi_z)}$$

n'_z = total pengamatan pada kelompok ke- z

o_z = banyaknya pengamatan kejadian sukses dalam kelompok ke- z

π_z = rata-rata estimasi probabilitas sukses kelompok ke- z

g = jumlah kelompok dengan tiap kelompok memuat $\frac{n}{10}$ pengamatan

$$\sum_{z=1}^g n'_z = n \quad ; z = 1, 2, \dots, g$$

Kriteria Uji

Tolak H_0 jika nilai $\hat{C} > X^2_{(\alpha; g-2)}$ dengan menggunakan $\alpha = 5\%$

3. Setelah pengujian parameter dan model, akan didapat model yang layak digunakan. Setelah itu dihitung nilai *odds ratio* dari masing-masing variabel prediktor. Harlan (2018) menjelaskan persamaan menghitung *odds ratio* adalah sebagai berikut:

$$OR_j = \exp(\beta_j)$$

4. Karena analisis regresi logistik dapat digunakan dalam pengklasifikasian, maka dapat dihitung nilai *apparent error rate* (APPER) untuk mengetahui seberapa besar kesalahan klasifikasi yang dilakukan oleh model regresi logistik *robust* tersebut. Rizal (2017) menjelaskan persamaan menghitung nilai APPER dan ketepatan klasifikasinya sebagai berikut:

$$APPER (\%) = \frac{\text{Jumlah prediksi salah}}{\text{Jumlah total prediksi}} \times 100\% \quad \text{Ketepatan Klasifikasi} = 100\% - APPER$$

5. Untuk mengetahui apakah hasil klasifikasi akurat, dilakukan pengujian keakuratan klasifikasi menggunakan nilai *Press's Q*. Rofiq (2016) menjelaskan pengujiannya sebagai berikut:

Hipotesis

H_0 : Hasil klasifikasi tidak akurat

H_1 : Hasil klasifikasi akurat

Statistik Uji

$$\text{Press's } Q = \frac{[N - n_i K]^2}{N(K-1)}$$

N = ukuran total sampel

n_i = jumlah pengamatan yang tepat diklasifikasikan ($n_{00} + n_{11}$)

K = jumlah kategori

Kriteria Uji

Tolak H_0 jika nilai *Press's Q* $> X^2_{(\alpha; 1)}$ dengan menggunakan $\alpha = 5\%$

3. Pembahasan dan Diskusi

Pemodelan Regresi Logistik Biner

Model regresi logistik dihitung mengikuti prosedur penelitian. Dengan menggunakan *software* RStudio, didapat model regresi logistik biner menggunakan metode MLE adalah sebagai berikut:

$$\hat{\pi}_i = \frac{\exp(-1,569552 - 0,3058617X_{i1} + 0,1862561X_{i2} + 0,1649086X_{i3} + 0,0675958X_{i4})}{1 + \exp(-1,569552 - 0,3058617X_{i1} + 0,1862561X_{i2} + 0,1649086X_{i3} + 0,0675958X_{i4})}$$

$\hat{\pi}_i$ = probabilitas terjadinya Demam Berdarah *Dengue* (DBD) pada pengamatan ke- i

X_{i1} = selisih kadar leukosit pada pengamatan ke- i

X_{i2} = selisih kadar trombosit pada pengamatan ke- i

X_{i3} = selisih kadar hematokrit pada pengamatan ke- i

X_{i4} = selisih kadar hemoglobin pada pengamatan ke- i

Pengujian Parameter dan Model Regresi Logistik *Robust*

Setelah mendapatkan model, dilakukan pengujian parameter secara simultan dan parsial, dan pengujian kelayakan model.

1. Uji Parameter Simultan dilakukan untuk mengetahui penduga parameter secara bersamaan apakah mempengaruhi variabel respon. Pengujiannya sebagai berikut:

Hipotesis

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$; Tidak ada variabel prediktor yang mempengaruhi variabel respon

$H_1: \min \beta_j \neq 0 ; j = 1, 2, 3, 4$; Minimal ada satu variabel prediktor yang mempengaruhi variabel respon

Statistik Uji

$$G = -2(-45,6229593 - (-38,0420127))$$

$$G = 15,1618932$$

Kriteria Uji

Nilai $G > X^2_{(0,05;4)}$ yaitu $15,1618932 > 9,49$. Maka Tolak H_0 .

Dari pengujian di atas didapat kesimpulan menolak hipotesis nol yang berarti bahwa minimal ada satu variabel prediktor yang berpengaruh terhadap variabel respon. Untuk mengetahui variabel prediktor mana yang berpengaruh, maka dilanjutkan dengan pengujian parameter secara parsial.

2. Uji Parameter Parsial dilakukan untuk mengetahui penduga parameter secara individu apakah mempengaruhi variabel respon. Pengujiannya sebagai berikut:

Hipotesis

$H_0: \beta_j = 0$; Variabel prediktor ke- j tidak berpengaruh terhadap variabel respon

$H_1: \min \beta_j \neq 0 ; j = 1, 2, 3, 4$; Variabel prediktor ke- j berpengaruh terhadap variabel respon

Statistik Uji

Statistik uji menggunakan uji Wald dengan nilainya pada tabel berikut:

Tabel 2. Statistik Uji Pengujian Parameter Parsial

Variabel	Estimasi	Std. Error	Wald
X_1	-0,30586	0,13120	-2,331
X_2	0,18626	0,07775	2,395
X_3	0,16491	0,18129	0,910
X_4	0,06760	0,50876	0,133

Kriteria Uji

Kriteria uji pada pengujian ini dapat dilihat pada tabel berikut:

Tabel 3. Kriteria Uji Pengujian Parameter Parsial

Variabel	Wald	$Z_{(0,025)}$	Keputusan
X_1	2,331	1,96	Tolak H_0
X_2	2,395	1,96	Tolak H_0
X_3	0,910	1,96	Terima H_0
X_4	0,133	1,96	Terima H_0

Dari pengujian di atas didapat pada variabel prediktor X_1 dan X_2 menolak hipotesis nol sementara X_3 dan X_4 menerima hipotesis nol. Sehingga dapat disimpulkan bahwa variabel X_1 dan X_2 berpengaruh secara signifikan terhadap variabel respon. Sementara variabel X_3 dan X_4 tidak berpengaruh terhadap variabel respon.

3. Uji Kelayakan Model dilakukan untuk mengetahui model yang didapat apakah layak digunakan. Pengujiannya sebagai berikut:

Hipotesis

H_0 : Model regresi logistik biner layak digunakan

H_1 : Model regresi logistik biner tidak layak digunakan

Statistik Uji

$$\hat{C} = 12,206$$

Kriteria Uji

Nilai $\hat{C} < X^2_{(0,05;7)}$ yaitu $12,206 < 15,51$. Maka Terima H_0 .

Dari pengujian di atas didapat kesimpulan menerima hipotesis nol yang berarti bahwa model regresi logistik biner yang didapat layak untuk digunakan.

Setelah melakukan pengujian parameter dan kelayakan model, didapat model regresi logistik biner dengan MLE yang dapat digunakan yaitu:

$$\hat{\pi}_i = \frac{\exp(-1,569552 - 0,3058617X_{i1} + 0,1862561X_{i2})}{1 + \exp(-1,569552 - 0,3058617X_{i1} + 0,1862561X_{i2})}$$

di mana

$\hat{\pi}_i$ = probabilitas terjadinya Demam Berdarah *Dengue* (DBD) pada pengamatan ke- i

X_{i1} = selisih kadar leukosit pada pengamatan ke- i

X_{i2} = selisih kadar trombosit pada pengamatan ke- i

Odds Ratio

Setelah melakukan pengujian parameter dan kelayakan model, didapat bahwa model yang layak digunakan dalam penelitian ini adalah model regresi logistik biner dengan variabel prediktor yang secara signifikan berpengaruh hanya variabel X_1 dan X_2 . Maka yang dihitung nilai *odds ratio* hanya parameter dari variabel X_1 dan X_2 .

Tabel 4. Odds Ratio Model Regresi Logistik Biner

Variabel	Estimasi	Odds Ratio
X_1	-0,30586	0,7365
X_2	0,18626	1,2047

Pada tabel di atas menunjukkan bahwa untuk variabel X_1 memiliki nilai *odds ratio* sebesar 0,7365 yang berarti jika pasien memiliki kenaikan satu satuan X_1 maka pasien memiliki resiko terkena DBD sebesar 0,7365 kali dibandingkan dengan pasien yang memiliki X_1 yang

tetap. Sementara untuk variabel X_2 memiliki nilai *odds ratio* sebesar 1,2047 yang berarti jika pasien memiliki kenaikan satu satuan X_2 maka pasien memiliki resiko terkena DBD sebesar 1,2047 kali dibandingkan dengan pasien yang memiliki X_2 yang tetap.

Nilai APPER dan Press's Q

Analisis regresi logistik dapat digunakan dalam pengklasifikasian. Dengan menggunakan model regresi logistik biner yang didapat, dilakukan klasifikasi variabel respon dengan nilai *cutoff* sebesar 0,5. Hasil klasifikasi dapat dilihat pada tabel berikut:

Tabel 5. Tabel Klasifikasi Respon Biner Model Regresi Logistik Biner

Observasi	Prediksi		Jumlah
	$Y_0 = 0$	$Y_1 = 1$	
$Y_0 = 0$	38	7	45
$Y_1 = 1$	13	12	25
Jumlah	51	19	70

Pada tabel di atas menunjukkan bahwa terdapat 38 pasien yang mengalami penyakit DD dan diprediksi mengalami penyakit DD dan 12 pasien yang mengalami penyakit DBD dan diprediksi mengalami penyakit DBD. Tingkat kesalahan klasifikasi dapat dilihat dengan menghitung nilai APPER sebagai berikut:

$$APPER (\%) = \frac{7 + 13}{70} \times 100\% = 28,57143 \%$$

$$\text{Ketepatan Klasifikasi} = 100 \% - 28,57143 \% = 71,42857 \%$$

Untuk mengetahui keakuratan dalam ketepatan klasifikasi, dilakukan dengan membandingkan nilai *Press's Q* dengan tabel *Chi-Square* berderajat bebas 1. Pengujiannya sebagai berikut:

Hipotesis

H_0 : Hasil klasifikasi model regresi logistik tidak akurat

H_1 : Hasil klasifikasi model regresi logistik akurat

Statistik Uji

Press's Q = 12,85714

Kriteria Uji

Nilai *Press's Q* > $\chi^2_{(0,05;1)}$ yaitu $12,85714 > 3,84$. Maka Tolak H_0 .

Dari pengujian di atas didapat kesimpulan menolak hipotesis nol yang berarti bahwa hasil klasifikasi model regresi logistik yang didapat dari hasil penelitian akurat.

4. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan hasil penelitian adalah model regresi logistik biner dengan menggunakan *maximum likelihood estimator* pada data penderita penyakit akibat infeksi virus *dengue* bulan Desember 2015 – Februari 2016 di RSU Haji Surabaya sebagai berikut:

$$\hat{\pi}_i = \frac{\exp(-1,569552 - 0,3058617X_{i1} + 0,1862561X_{i2})}{1 + \exp(-1,569552 - 0,3058617X_{i1} + 0,1862561X_{i2})}$$

dengan $\hat{\pi}_i$ adalah probabilitas terjadinya Demam Berdarah *Dengue* (DBD) pada pengamatan ke- i , X_{i1} adalah selisih kadar leukosit pada pengamatan ke- i pada pemeriksaan hari pertama dan hari ketiga, dan X_{i2} adalah selisih kadar trombosit pada pengamatan ke- i pada pemeriksaan hari pertama dan hari ketiga.

Acknowledge

Peneliti mengucapkan terima kasih kepada Allah SWT, Bapak Sofyan Mawardi & Ibu Neneng Suningsih, Keluarga Kurniaatmadja & M. Sadeli, Dosen dan Staf Prodi Statistika Unisba, Mahasiswa Statistika Unisba Angkatan 2017, dan semua pihak yang terlibat dalam penelitian. Semoga Allah SWT yang akan membalas kebbaikannya.

Daftar Pustaka

- [1] Apriliyanti, F. 2021. Perbandingan Ketepatan Klasifikasi Analisis Regresi Model Logit Biner dan Model Probit Biner terhadap Faktor-Faktor yang Mempengaruhi Risiko Penyakit Jantung Koroner. *Prosiding Statistika*, Vol. 7 No.1
- [2] Hajarisman, N. 2009. *Analisis Data Kategorik*. Bandung: Prodi Statistika Universitas Islam Bandung
- [3] Harlan, J. 2018. *Analisis Regresi Logistik*. Depok: Gunadarma.
- [4] Kemenkes. 2017. *Pedoman Pencegahan dan Pengendalian Demam Berdarah Dengue di Indonesia*. Jakarta: Kementerian Kesehatan RI.
- [5] Kotimah, M. K., & Wulandari, S. P. 2014. Model Regresi Logistik Biner Stratifikasi Pada Partisipasi Ekonomi Perempuan Di Provinsi Jawa Timur. *Jurnal Sains Dan Seni Pomits*, 2337-3520.
- [6] Nugraha, J. 2013. *Pengantar Analisis Regresi Logistik*. Yogyakarta: Deepublish.
- [7] Rizal, S. 2017. *Bootstrap Aggregating Multivariate Adaptive Regression Spline (Bagging Mars) Untuk Klasifikasi Indeks Pembangunan Manusia Provinsi Jawa Tengah*. Semarang: Universitas Muhamadiyah Semarang.
- [8] Rofiq, A. 2016. Perbandingan Analisis Diskriminan Fisher dan Naive Bayes Untuk Klasifikasi Risiko Kredit (Studi Kasus Debitur di Koperasi Jateng Amanah Mandiri Cabang Sukorejo Kendal). *Journal Gaussian*, 1-10.
- [9] Sari, W. N. 2017. *Pemodelan Regresi Logit Biner Robust Pada Data Klasifikasi Penderita Penyakit Akibat Infeksi Virus Dengue Berdasarkan Penduga Bianco Yohai*. Surabaya: Universitas Airlangga.
- [10] Wulandari, R. 2018. *Penerapan Model Regresi Logistik Robust Pada Status Hiperkolesterolemia di RSUD DR. Moewardi Surakarta Menggunakan Estimasi Bianco-Yohai*. Surakarta: Universitas Sebelas Maret.
- [11] Utama Muhammad Bangkit Riksa, Hajarisman Nusar. (2021). *Metode Pemilihan Variabel pada Model Regresi Poisson Menggunakan Metode Nordberg*. *Jurnal Riset Statistika*, 1(1), 35-42.