

**LAPORAN KEMAJUAN
PENELITIAN DASAR
UNIVERSITAS LAMPUNG**



**MENANGANI MULTIKOLINEARITAS DALAM ANALISIS REGRESI
LOGISTIK: STUDI PERBANDINGAN**

TIM PENGUSUL :

1. Prof. Ir. Netti Herawati, M.Sc., Ph.D. (6169478)
2. Dr. Khoirin Nisa, M. (6050683)
3. Subian Saidi, S.Si., M.Si. (6681591)

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG**

2024

HALAMAN PENGESAHAN
LAPORAN KEMAJUAN PENELITIAN DASAR UNIVERSITAS LAMPUNG

Judul : Menangani Multikolinearitas Dalam Analisis Regresi Logistik: Studi Perbandingan
Fokus Riset : IPTEK
Manfaat saintifik/sosial : Menambah keilmuan dalam penggunaan metode statistik dan aplikasi statistika pada data real

Ketua Pengabdian:
a. Nama Lengkap : Prof. Ir. Netti Herawati, M.Sc., Ph.D.
b. SINTA ID : 6169478
c. Jabatan fungsional : Guru Besar
d. Program studi : Matematika
e. Nomor HP : 081273809624
f. Alamat surel (e-mail) : netti.herawati@fmipa.unila.ac.id

Anggota Pengabdian (1)
a. Nama Lengkap : Dr. Khoirin Nisa, M.Si.
b. NIDN : 0026077401
c. SINTA ID : 6050683
d. Jabatan fungsional : Lektor
e. Program studi : Matematika
f. Alamat surel (e-mail) : nisa.mahfudh@gmail.com


Anggota Pengabdian (2)
g. Nama Lengkap : Subian Saidi, S.Si. M.Si.
h. SINTA ID : 6681591
i. Jabatan fungsional : Asisten Ahli
j. Program studi : Matematika
k. Alamat surel (e-mail) : subian.saidi@fmipa.unila.ac.id

Mahasiswa yang terlibat (1)
a. Nama Lengkap : Eva Selviana
b. NPM : 1817031075
c. Program studi : Matematika
d. Alamat surel (e-mail) : evaselviana@gmail.com
Jumlah alumni yang terlibat : 1 orang
Jumlah staf yang terlibat : 1 orang
Mitra penelitian : 1 orang
Lokasi penelitian : Bandar Lampung
Lama penelitian : 6 bulan
Biaya penelitian : Rp. 25.000.000,-
Sumber dana : DIPA BLU Unila Tahun 2024

Wakil Dekan Bidang Akademik dan
Kerjasama FMIPA

Mulyono, S.Si., M.Si., Ph.D.
NIP. 197306112000031000

Bandar Lampung, 7 Juni 2024
Ketua Peneliti,


Prof. Ir. Netti Herawati, M.Sc., Ph.D.
NIP. 196501251990032001

Ketua LPPM Universitas Lampung


Dr. Eng. Ir. Dikpride D. S. W.T., IPM, ASEAN Eng.
NIP. 197207281998032001

Menangani Multikolinearitas dalam Analisis Regresi Logistik: Studi Perbandingan

RINGKASAN

Metode Maximum Likelihood Estimation (MLE) merupakan metode yang sering digunakan pada regresi logistik. Analisis regresi logistik yang merupakan metode statistika untuk menganalisis hubungan antara variabel dependen yang bersifat dikotomik dan beberapa variabel independen. Dalam beberapa penelitian regresi logistik, sering ditemukan permasalahan dimana data mengandung multikolinearitas. Namun, metode MLE tidak akan stabil digunakan ketika data yang diteliti mengandung multikolinearitas. Beberapa metode tersedia untuk mengatasi multikolinearitas pada regresi logistik misalnya LASSO dan Liu. Kelebihan metode Liu dibanding metode lainnya yaitu memiliki nilai tetapan yang merupakan fungsi linear, sehingga metode Liu mudah ditentukan. Penelitian ini bertujuan untuk mengetahui perbandingan performa estimator *Least Absolute Shrinkage and Selection Operator* (LASSO), Liu dan metode MLE dalam mengatasi multikolinearitas melalui simulasi Monte Carlo dengan $n = 25, 50, \text{ dan } 75$ serta 6 variabel independen dengan tingkat multikolinearitas sebesar 0,99 pada data simulasi dengan pengulangan sebanyak 1000 kali. Metode terbaik dibandingkan berdasarkan nilai MSE terkecil.

Metode yang kami kembangkan akan memberikan kontribusi besar bagi pengembangan ilmu statistika pada umumnya, dan bagi pengembangan teknik statistika dalam mengatasi multikolinearitas pada khususnya.

I. PENDAHULUAN

1.1 Latar Belakang

Untuk mengetahui hubungan antara satu atau lebih variabel bebas terhadap variabel tak bebas Analisis regresi merupakan salah satu metode statistika. Umumnya, data yang digunakan pada analisis regresi adalah data kontinu berdistribusi normal. Namun, dalam beberapa penelitian, variabel tak bebas (dependen) yang diteliti menggunakan data kategorik, yang menyatakan kejadian sukses dan gagal. Dalam analisis regresi, model yang digunakan dalam menganalisis hubungan antara variabel bebas dengan variabel tak bebas yang menggunakan data bersifat dikotomik ialah model regresi logistik. Menurut Hosmer & Lemeshow (2000), Analisis regresi logistik merupakan salah satu metode statistika yang sering digunakan dalam menganalisa data respon yang bersifat biner atau multinomial.

Salah satu metode untuk mencari estimasi parameter logistik adalah metode *Maximum Likelihood Estimation* (MLE). Ide dasar dari metode ini adalah menggunakan nilai suatu ruang parameter yang menghubungkan data observasi yang memiliki kemungkinan (*likelihood*) terbesar sebagai penduga dari parameter yang tidak diketahui. Prosedur dari metode *Maximum Likelihood Estimation* (MLE), yaitu dengan cara memaksimumkan nilai $L(\theta)$ atau disebut dengan *conditional log-likelihood function*.. Metode ini baik digunakan bila asumsi multikolinearitas terpenuhi. Multikolinearitas dapat berpengaruh pada tingkat akurasi prediksi model dan menyebabkan kesalahan dalam pengambilan keputusan (Montgomery and Peck, 1992; Drapper and Smith, 1998). Namun, dalam beberapa penelitian regresi logistik, sering ditemukan permasalahan dimana data yang digunakan mengandung multikolinearitas.

Adanya multikolinearitas akan menyebabkan estimasi parameter regresi menjadi tidak akurat yang mengakibatkan nilai rata-rata kuadrat error menjadi besar. Menurut Myers (1990), terjadinya multikolinearitas diantara variabel-variabel bebas dapat mengakibatkan konsekuensi penting bagi penafsiran dan penggunaan model regresi

dugaan, karena dapat menyebabkan tanda dari koefisien regresi menjadi salah atau keputusan menjadi tidak signifikan.

Metode MLE yang digunakan untuk mengestimasi parameter pada regresi logistik tidak akan stabil digunakan ketika data yang diteliti mengandung multikolinearitas. Beberapa metode dikembangkan untuk dapat mengatasi permasalahan multikolinearitas, diantaranya yaitu metode *Ridge*, *LASSO*, *Elastic Net*, dan Liu. Juga telah dilakukan penelitian tentang pengaruh multikolinearitas pada regresi berganda dan regresi logistik (Herawati *et al.* 2018; Hastie, *et al.*, 2008; Toka, 2016). Hoerl & Kennard (1970) memperkenalkan metode Ridge dimana analisis regresi ridge dianggap memiliki sifat optimal dalam menangani kasus multikolinearitas. Liu (1993) menyarankan penduga lain yang merupakan kombinasi dari metode Stein dan metode Ridge, dimana metode tersebut memiliki sifat optimal seperti metode Ridge. Namun kelebihan metode Liu dibandingkan metode Ridge yaitu nilai tetapan pada metode Liu merupakan fungsi linear sehingga metode Liu mudah ditentukan (Qasim, dkk., 2019).

Berdasar pada penjelasan di atas, performa metode LASSO dan Liu akan diteliti menggunakan data simulasi pada model regresi logistik yang mengandung multikolinearitas. Kemudian akan dibandingkan dengan metode *Maximum Likelihood Estimation* (MLE) berdasarkan nilai MSE (*Mean Square Error*) dari masing-masing metode.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah Meneliti efisiensi metode *Least Absolute Shrinkage and Selection Operator* (LASSO) dan Liu untuk menangani masalah multikolinearitas pada regresi logistik.

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah

1. Mengintegrasikan hasil penelitian mengenai perbandingan metode MLE dengan metode *Least Absolute Shrinkage and Selection Operator* (LASSO), dan *Lu* pada model regresi logistik biner pada data simulasi dan data real berdasarkan nilai MSE. Pada pembelajaran mata kuliah analisis regresi terapan.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah memberikan informasi dan masukan kepada peneliti maupun pembaca mengenai performa metode MLE, LASSO dan Liu untuk menganalisis data yang mengandung multikolinieritas pada regresi logistik serta menambah pengetahuan bagi penulis dan mahasiswa dalam penggunaan analisis regresi terapan.

I. TINJAUAN PUSTAKA

2.1 Regresi Logistik

Analisis regresi logistik merupakan analisis regresi di mana variabel tak bebas memiliki sifat biner atau dikotomis dengan satu atau lebih variabel bebas (Hosmer & Lemeshow, 2000; Myers, 1990). Pada regresi logistik, variabel tak bebas atau *dependent* berskala kategorik. Variabel tak bebas yang dinotasikan dengan y bersifat biner atau *dikotomis* yang mempunyai dua nilai yaitu 0 dan 1. Dengan demikian, variabel y mengikuti distribusi Bernoulli untuk setiap observasi tunggal. Fungsi probabilitas untuk setiap observasi diberikan sebagai berikut:

$$f(y) = \pi^y(1 - \pi)^{1-y} \quad y = 0,1 \quad (2.1)$$

Model umum $\pi(x)$ dinotasikan sebagai:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)} \quad (2.2)$$

Persamaan (2.2) disebut fungsi regresi logistik yang menunjukkan hubungan antara variabel bebas dan probabilitas yang tidak linier, sehingga untuk mendapatkan hubungan yang linier dilakukan transformasi yang sering disebut dengan transformasi logit. Bentuk logit dari $\pi(x)$ dinyatakan sebagai $g(x)$, yaitu:

$$\text{logit}[\pi(x)] = g(x) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (2.3)$$

Menurut Hosmer & Lemeshow (2000), persamaan (2.3) merupakan fungsi regresi logistik yang disebut model regresi logistik berganda. Asumsi Analisis Regresi Logistik antara lain tidak ada multikolinieritas antar variabel independen.

2.2 Multikolinieritas

Multikolinieritas terjadi bila terdapat dua atau lebih peubah bebas yang saling berkaitan. Multikolinieritas dapat dideteksi menggunakan nilai *Variance Inflation Factor* (VIF) (Montgomery & Peck, 2012). Nilai VIF dapat dicari menggunakan rumus $VIF_{(j)} = \frac{1}{(1-R_j^2)}$ merupakan koefisien determinasi yang didapat dari

variabel prediktor X_j yang diregresikan dengan variabel prediktor lainnya. Jika X_j tidak berkorelasi dengan peubah bebas lain, maka R_j^2 akan bernilai kecil dan nilai VIF akan mendekati 1. Sebaliknya jika X_j mempunyai korelasi dengan peubah bebas lain, maka R_j^2 akan mendekati 1 dan nilai VIF menjadi besar. Jika nilai VIF lebih besar dari 10, maka menunjukkan adanya multikolinearitas (Draper & Smith, 1992).

2.3 Maximum Likelihood Estimation (MLE)

Maximum Likelihood Estimation (MLE) merupakan metode estimasi parameter dalam regresi logistik dengan distribusi variabel yang telah diketahui. Metode tersebut mengestimasi parameter β dengan cara memaksimalkan fungsi *likelihood* dan mensyaratkan bahwa data harus mengikuti suatu distribusi tertentu. Ketika y_i menyebar binomial dengan i banyaknya sampel maka fungsi kepekatan peluang bersama adalah :

$$f(y|\beta) = \prod_{i=1}^N \frac{n_i!}{y_i!(n_i - y_i)!} \pi_i^{y_i} (1 - \pi_i)^{n_i - y_i} \quad (2.4)$$

Dengan demikian didapatkan fungsi *likelihood*:

$$L(\beta|y) = \prod_{i=1}^N \frac{n_i!}{y_i!(n_i - y_i)!} \pi_i^{y_i} (1 - \pi_i)^{n_i - y_i} \quad (2.5)$$

MLE untuk menduga parameter β yaitu dengan memaksimalkan nilai fungsi *likelihood* $L(\beta) = \prod_{i=1}^N y_i \left(\sum_{p=0}^P x_{ip} \beta_p \right) - n_i \cdot \log \left(1 + \exp \left(\sum_{p=0}^P x_{ip} \beta_p \right) \right)$

Pendugaan kemungkinan maksimum untuk β adalah $\hat{\beta}_{MLE} = (\mathbf{X}^T \widehat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}^T \widehat{\mathbf{W}} \mathbf{z}$ dimana \mathbf{z} merupakan vektor kolom dengan i elemen sama dengan $\text{logit}(\hat{\pi}_i) + \frac{y_i \hat{\pi}_i}{\hat{\pi}_i (1 - \hat{\pi}_i)}$ dan $\widehat{\mathbf{W}} = \text{diag}[\hat{\pi}_i (1 - \hat{\pi}_i)]$ yang merupakan penduga tak bias dari β (Mansson et al., 2012).

2.3 Least Absolute Shrinkage and Selection Operator (LASSO)

Metode Least Absolute Shrinkage and Selection Operator (LASSO) dapat digunakan untuk mengatasi masalah multikolinearitas. LASSO diperkenalkan oleh Tibshirani (1996), dimana LASSO bekerja menyusutkan koefisien (parameter β) yang berkorelasi menjadi tepat pada nol atau mendekati nol. Menurut Hastie, et al. (2015), pendugaan parameter pada LASSO adalah sebagai berikut:

$$\pi_i = \frac{1}{1 + \exp(-\beta^T x_i)} \quad (2.6)$$

Kendala Lagrangian (L^1 -norm) dapat digabungkan dalam pendugaan parameter *log-likelihood* pada regresi logistik. Persamaan gabungan *log-likelihood* sebagai untuk vektor β adalah:

$$\begin{aligned} l(\beta) &= L(\beta \mid y_1, \dots, y_n) \\ &= \sum_{i=1}^n \left[y_i \ln \left(\frac{1}{1 + \exp(-\beta^T x_i)} \right) + (1 - y_i) \left(\frac{\exp(\beta^T x_i)}{1 + \exp(\beta^T x_i)} \right) \right] \\ &= \sum_{i=1}^n [(1 - y_i)\beta^T x_i + \ln(1 + \exp(-\beta^T x_i))] \end{aligned} \quad (2.7)$$

Persamaan gabungan antara log-likelihood dengan kendala Lagrangian menghasilkan persamaan sebagai berikut.

$$l(\beta) = - \sum_{i=1}^n [(1 - y_i)\beta^T x_i + \ln(1 + \exp(-\beta^T x_i))] - \lambda \sum_{j=1}^p |\beta_j| \quad (2.8)$$

Sehingga kita peroleh peduga parameter regresi logistik dengan LASSO

$$\beta_{\lambda}^{LASSO} = \operatorname{argmax} \{l(\beta) - \lambda \sum_{j=1}^p |\beta_j|\} \quad (2.9)$$

λ merupakan nilai bias dalam metode LASSO, dimana nilai $\lambda > 0$. Nilai λ diperoleh dengan beberapa metode, diantaranya metode validasi silang (Tibshirani, 1996; Fonti & Belitser, 2017).

2.4 Metode Liu

Metode Liu merupakan metode alternatif regresi logistik berupa estimator susut bias dan penduga langsung generalisasi yang diusulkan untuk model regresi linier oleh Liu (1993) guna mengatasi masalah multikolinearitas. Estimator Liu memiliki

keunggulan dibanding dengan metode yang lain, seperti memiliki nilai *Scalar Mean Square Error* (SMSE) yang lebih kecil dibandingkan dengan taksiran *Ridge*. Oleh karena itu, Liu (1993) menyarankan estimator lain dimana parameter yang diperoleh dari estimator ini memiliki manfaat sebagai fungsi linier dari parameter penyusutan d .

Penyusutan parameter d dapat mengambil nilai antara nol dan satu dan ketika d kurang dari satu maka kita memiliki $\|\widehat{\beta}_d\| \leq \|\widehat{\beta}_{MLE}\|$. Metode Liu selanjutnya dikembangkan oleh Mansson, dkk (2012) pada model regresi logistik sebagai:

$$\widehat{\beta}_d = (X'\widehat{W}X + I)^{-1}(X'\widehat{W}X + dI)\widehat{\beta}_{MLE} \quad (2.10)$$

Adapun penaksiran untuk nilai d yang diusulkan oleh Hoerl & Kennard (1970) yaitu:

$$d_1 = \max \left[0, \frac{\widehat{a}_{j\max}^2 - 1}{\frac{1}{\lambda_{j\max}} + \widehat{a}_{j\max}^2} \right] \quad (2.11)$$

Selanjutnya, estimator berikut yang didasarkan pada ide-ide dalam Kibria (2003), diusulkan:

$$d_2 = \max \left[0, \text{median} \frac{\widehat{a}_j^2 - 1}{\frac{1}{\lambda_j} + \widehat{a}_j^2} \right] \quad (2.12)$$

$$d_3 = \max \left[0, \frac{1}{p} \sum_j \left(\frac{\widehat{a}_j^2 - 1}{\frac{1}{\lambda_j} + \widehat{a}_j^2} \right) \right] \quad (2.13)$$

Akhirnya, estimator berikut diusulkan dimana pada estimator ini kuantil lain selain median digunakan dan berhasil diterapkan oleh Khalaf dan Shukur (2005).

$$d_4 = \max \left[0, \max \frac{\widehat{a}_j^2 - 1}{\frac{1}{\lambda_j} + \widehat{a}_j^2} \right] \quad (2.14)$$

$$d_5 = \max \left[0, \min \frac{\widehat{a}_j^2 - 1}{\frac{1}{\lambda_j} + \widehat{a}_j^2} \right] \quad (2.15)$$

$$d_6 = \max \left[0, \text{median} \left[\frac{\hat{a}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} - \hat{a}_j^2} \right] \right] \quad (2.16)$$

2.5 Pengukuran Perbandingan Metode

Untuk mengukur metode mana yang mempunyai keakuratan tinggi dan paling baik dalam memodelkan data, *Mean Square Error* (MSE) digunakan. MSE adalah salah satu pengukuran kesalahan yang terkenal dan mudah digunakan (Ghozali, 2006). Semakin kecil nilai MSE, maka semakin akurat nilai suatu pemodelan. Pada permasalahan multikolinearitas, metode terbaik merupakan metode yang dapat digunakan untuk memperbaiki multikolinearitas. Adapun perbaikan masalah multikolinearitas ini akan dilihat berdasarkan rata-rata dari *Mean Square Error* (MSE) dari hasil estimasi parameter β sebagai berikut:

$$\text{MSE } \hat{\beta} = \frac{1}{m} \sum_{j=1}^m (\hat{\beta}_j - \beta_i)^2 \quad ; j = 1, 2, \dots, m \quad (2.17)$$

Dengan $\hat{\beta}_j$ = Penduga parameter regresi; β_i = Parameter regresi dan m = Banyaknya ulangan

Selain itu juga digunakan metode AIC yang merupakan salah satu metode yang digunakan untuk memilih model regresi terbaik. Metode AIC ditemukan oleh Akaike (Grasa, 1989). Metode AIC didasarkan pada metode MLE. Menurut metode AIC, model regresi terbaik adalah model yang memiliki nilai AIC terkecil (Widarjono, 2007). Adapun rumus AIC sebagai berikut.

$$AIC = e^{\frac{2k}{n} \frac{\sum_{i=1}^n \hat{u}_i^2}{n}} \quad (2.18)$$

Dengan k = Jumlah parameter yang diestimasi dalam model regresi; n = Jumlah observasi dan u = residual /sisa.

Sedangkan untuk digunakan untuk mengetahui signifikan atau tidaknya tiap-tiap parameter variabel bebas terhadap variabel terikat digunakan uji statistik Wald. Pengujian hipotesis uji Wald sebagai berikut.

$$H_0 : \hat{\beta}_i = 0$$

$$H_1 : \hat{\beta}_i \neq 0$$

Dengan statistik Uji Wald: $W_i = \left[\frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \right]^2 ; i = 0,1,2, \dots, p$

Statistik uji Wald berdistribusi Chi-Square dengan derajat bebas 1 (Haloho, dkk., 2013). Kriteria uji wald adalah tolak H_0 jika $W_i > X_{1,a}^2$ atau tidak tolak H_0 jika $W < X_{1,a}^2$ dengan $a = 0.05$ dan $df = 1$.

II. METODE PENELITIAN

3.1 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data simulasi dan data real yang mengandung multikolinearitas dan tidak mengandung multikolinearitas. Untuk data simulasi yakni data yang dibangkitkan dengan variabel bebas sebanyak $p=6$ dengan $n=25, 50, 75$ yang diulang sebanyak 1000 kali dengan menggunakan *software* R versi 3.6.1. Untuk mendapatkan data multikolinearitas pada setiap himpunan data X_p dibangkitkan menggunakan simulasi Monte Carlo dengan persamaan $X_p = \sqrt{(1 - \rho^2)}z_{ij} + \rho z_{i(p+1)}$; $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, 6$ dimana $z_{ij} \sim N(0,1)$ dengan $\rho = 0,99$. Lalu dari nilai X_p yang telah dibangkitkan maka akan didapatkan variabel terikat (Y) yang dibangkitkan dengan distribusi binomial dengan probabilitas regresi logistik biner yaitu $Y = \pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}$ dengan $\beta_0 = 0$ dan $\beta_1 = \beta_2 = \dots = \beta_p = 1$ serta sebaran binomial($n, 1, \text{prob} = \pi(x)$).

3.2 Metode Penelitian

Secara rinci langkah-langkah penelitian yang dilakukan adalah sebagai berikut:

1. Melakukan simulasi data regresi logistik.
2. Mengidentifikasi multikolinearitas dengan melihat nilai VIF pada data simulasi dan data real.
3. Menentukan koefisien parameter (β) dengan analisis regresi logistik menggunakan *Maximum Likelihood Estimation* pada data simulasi dan data real.
4. Melakukan pendugaan parameter dengan metode MLE, LASSO, dan Liu pada data simulasi
5. Melakukan perhitungan MSE pada metode MLE, LASSO, dan Liu data simulasi
6. Menarik kesimpulan

Tabel Pembagian Tugas

No.	Posisi	Peran/Tanggung Jawab
1	Ketua	Melakukan setiap proses penelitian, Mengarahkan anggota penelitian mulai dari studi pustaka, analisis data sampai dengan diperoleh luaran
2	Anggota (1)	Membantu setiap proses penelitian mulai dari melakukan studi pustaka dan analisis data
3	Anggota (2)	Membantu setiap proses penelitian mulai dari melakukan studi pustaka dan analisis data
4	Mahasiswa	Membantu dalam mencari studi pustaka

IV. HASIL DAN PEMBAHASAN

Setelah data dibangkitkan, dilakukan pengecekan nilai korelasi antarvariabel dan VIF masing-masing variabel bebas. Adapun nilai VIF ditunjukkan sebagai berikut.

Table 1. Nilai VIF

<i>n</i>	Variabel Independen					
	X_1	X_2	X_3	X_4	X_5	X_6
25	60.1037	86.6308	44.8282	81.0369	104.1233	72.4655
50	50.6109	44.7655	67.2899	41.7368	42.2690	43.5353
75	61.2666	79.4079	52.0325	61.8288	69.9952	47.1938

Berdasarkan Tabel 1, nilai VIF > 10 diperoleh untuk semua variabel independen di semua data. Hal ini berarti terjadi multikolinearitas karena variabel bebas terlihat saling berkorelasi. Selain VIF, ditunjukkan nilai korelasi antarvariabel untuk $n=25,50,75$.

Tabel 2. Nilai korelasi antarvariabel untuk $n=25$

$n = 25$	X1	X2	X3	X4	X5	X6
X1	1	0.9855553	0.9813981	0.9892030	0.9863767	0.9820822
X2	0.9855553	1	0.9847641	0.9876933	0.9917750	0.9873204
X3	0.9813981	0.9847641	1	0.9841272	0.9822016	0.9738643
X4	0.9892030	0.9876933	0.9841272	1	0.9859681	0.9864211
X5	0.9863767	0.9917750	0.9822016	0.9859681	1	0.9908091
X6	0.9820822	0.9873204	0.9738643	0.9864211	0.9908091	1

Tabel 3. Nilai korelasi antarvariabel untuk n=50

n = 50	X1	X2	X3	X4	X5	X6
X1	1	0.9830985	0.9855550	0.9785363	0.9817809	0.9818400
X2	0.9830985	1	0.9819366	0.9799110	0.9799350	0.9807683
X3	0.9855550	0.9819366	1	0.9857661	0.9831067	0.9835635
X4	0.9785363	0.9799119	0.9857661	1	0.9765319	0.9754445
X5	0.9817809	0.9799350	0.9831067	0.9765319	1	0.9814711
X6	0.9818400	0.9807683	0.9835635	0.9754445	0.9814711	1

Tabel 4. Nilai korelasi antarvariabel untuk n=75

n = 75	X1	X2	X3	X4	X5	X6
X1	1	0.9873363	0.9832745	0.9878074	0.9859102	0.9841967
X2	0.9873362	1	0.9873720	0.9882705	0.9890132	0.9852008
X3	0.9832745	0.9873720	1	0.9818524	0.9866369	0.9821433
X4	0.9878074	0.9882705	0.9818524	1	0.9868148	0.9821531
X5	0.9859102	0.9890132	0.9866369	0.9868148	1	0.9853063
X6	0.9841967	0.9852008	0.9821433	0.9821531	0.9853063	1

Berdasarkan Tabel 2-4, dapat dilihat bahwa antarvariabel bebas memiliki nilai korelasi yang tinggi diatas 0,5. Hal ini menunjukkan bahwa terdapat hubungan yang tinggi antarvariabel dan memungkinkan adanya multikolinearitas antarvariabel. Setelah nilai korelasi antarvariabel dan VIF telah dilakukan, selanjutnya adalah analisis menggunakan metode MLE,

LASSO, dan Liu sehingga diperoleh nilai duga $\beta_1 - \beta_6$ pada metode MLE, LASSO, dan Liu. Setelah didapat nilai duga, selanjutnya menentukan nilai *Standard Error* (SE) dari masing-masing penduga.

Table 5. Nilai $\hat{\beta}_p$ dan SE pada metode MLE, LASSO, dan Liu untuk n=25

n = 25	$\hat{\beta}_p$			SE		
	MLE	LASSO	LIU	MLE	LASSO	LIU
$\hat{\beta}_1$	1.15e+14	1.1168	-1.24e - 05	19.389348	0.9573	0.0999
$\hat{\beta}_2$	-1.07e+12	0.6428	-3.16e - 07	8.15527051	4.4381	0.0999
$\hat{\beta}_3$	3.28ec+ 13	2.0811	-2.32e - 06	19.4251196	0.0157	0.0999
$\hat{\beta}_4$	-1.40e+14	0.6656	-1.24e - 05	5.78962215	2.884	0.0996
$\hat{\beta}_5$	-6.59e+12	0.8240	-2.28e - 08	37.2348995	0.1000	0.0999
$\hat{\beta}_6$	1.002e+14	2.8538	-1.23e - 08	15.3285868	1.4645	0.1000

Tabel 6. Nilai $\hat{\beta}_p$ dan SE pada metode MLE, LASSO, dan Liu untuk n=50

n = 50	$\hat{\beta}_p$			SE		
	MLE	LASSO	LIU	MLE	LASSO	LIU
$\hat{\beta}_1$	125.5341	0.4807	0.0017	16.5464	45.0760	0.09999014
$\hat{\beta}_2$	64.3883	0.7259	0.0001	4.7514	3.4023	0.0999736

$\hat{\beta}_3$	276.4802	1.1891	0.0004	45.0222	156.6259	0.09988864
$\hat{\beta}_4$	2.9544	1.6238	0.0018	5.6277	11.7517	0.09999941
$\hat{\beta}_5$	-145.624	2.1738	5.12e – 05	15.0407	0.0001	0.09999941
$\hat{\beta}_6$	82.0230	2.1177	1.50e – 05	3.6686	176.3854	0.1

Tabel 7. Nilai $\hat{\beta}_p$ dan SE pada metode MLE, LASSO, dan Liu untuk n=75

n = 75	$\hat{\beta}_p$			SE		
	MLE	LASSO	LIU	MLE	LASSO	LIU
$\hat{\beta}_1$	3.2821	0.9562	0.0015	0.71536876	17.9985	0.09999991
$\hat{\beta}_2$	18.9339	1.1492	0.0003	3.86477248	0.1000	0.09999903
$\hat{\beta}_3$	4.3312	1.3643	0.0005	12.0116452	16.6810	0.09999583
$\hat{\beta}_4$	16.9225	1.7549	0.0016	5.92102649	1.7256	0.1000
$\hat{\beta}_5$	6.7331	0.9265	5.05e – 06	3.12660553	3.1920	0.1000
$\hat{\beta}_6$	12.6864	1.0088	3.72e – 06	5.06832898	4.8707	0.1000

Berdasarkan Tabel 5-7, dapat disimpulkan bahwa penduga MLE, LASSO, dan Liu pada n=25,50,75 memiliki nilai $\hat{\beta}_p$ yang bervariasi. Nilai duga $\hat{\beta}_p$ pada metode LASSO dan Liu, lebih mendekati ke parameter sebenarnya yaitu $\hat{\beta}_p=1$, dibandingkan pada metode MLE.

Sedangkan untuk nilai SE, SE pada metode Liu lebih kecil dibandingkan SE pada metode MLE dan LASSO. Hal ini menunjukkan bahwa pendugaan parameter untuk $\hat{\beta}_1 - \hat{\beta}_6$ menggunakan metode Liu jauh lebih baik dibandingkan

metode MLE dan LASSO. Semakin kecil nilai suatu SE suatu parameter, maka semakin baik dalam menduga parameter dari data yang mengandung multikolinearitas, sehingga penduga parameter baik digunakan dalam model yang digunakan.

Setelahnya dilakukan perbandingan nilai MSE (*Mean Square Error*) pada metode MLE, LASSO, dan Liu untuk menentukan metode terbaik.

Tabel 8. Nilai MSE dan AIC pada MLE, LASSO, LIU

Ukuran sampel	MSE			AIC		
	MLE	LASSO	LIU	MLE	LASSO	LIU
n=25	267.36	1.3455	0.0599	14.05	-48.2964	-59.4711
n=50	349.7437	1.3903	0.0599	14.03	-105.0664	-109.4559
n=75	343.0046	1.2709	0.0598	14.00	-172.44	-177.8635

Berdasarkan Tabel 8, dapat dilihat bahwa pada simulasi data n=25,50,75, nilai MSE pada metode Liu lebih kecil dibanding dengan MSE pada metode MLE dan LASSO. Nilai MSE metode Liu rata-rata 0.0599 untuk ketiga sampel data. Sehingga dapat disimpulkan bahwa metode terbaik yang dapat digunakan untuk mengatasi multikolinearitas adalah metode Liu. Hal ini dikarenakan, semakin kecil nilai MSE suatu model, maka nilai pemodelan yang didapat akan semakin baik dan akurat. Selain itu, nilai AIC pada metode Liu juga terbukti lebih kecil dibandingkan MLE dan LASSO. Semakin besar jumlah sampel, semakin kecil nilai AIC yang didapatkan. Sehingga, dapat disimpulkan bahwa metode terbaik yang dapat digunakan untuk mengatasi multikolinearitas berdasarkan MSE dan AIC adalah metode Liu.

1. Penerapan Metode Liu pada Data Asli

Data real yang digunakan pada penelitian ini berupa data sekunder yang diperoleh dari Badan Pusat Statistik (BPS). Data yang digunakan yaitu persentase penduduk miskin di Indonesia tahun 2018 (Y) dengan faktor yang mempengaruhi yaitu kepadatan penduduk, Indeks Pembangunan Manusia (IPM), rata-rata lama sekolah, pengeluaran per kapita, angka melek huruf, dan angka harapan hidup. Menurut BPS, rata-rata persentase penduduk miskin di Indonesia tahun 2018 sebesar 9,82%. Dimana persentase ini dikelompokkan menjadi 2, kategori miskin ($> 9,82\%$) dan kategori tidak miskin ($\leq 9,82\%$).

Tabel 9. Data Persentase penduduk miskin tahun 2018

NO	WILAYAH	Y	X1	X2	X3	X4	X5	X6
1	ACEH	1	91.13	71.19	9.09	9186	98.03	69.64
2	SUMATERA UTARA	0	197.52	71.18	9.34	10391	99.07	68.61
3	SUMATERA BARAT	0	128.11	71.73	8.76	10638	99.07	69.01
4	RIAU	0	78.31	72.44	8.92	10968	99.2	71.19
5	JAMBI	0	71.32	70.65	8.23	10357	98.15	70.89
6	SUMATERA SELATAN	1	91.39	69.39	8	10652	98.66	69.41
7	BENGKULU	1	98.56	70.64	8.61	10162	97.91	68.84
8	LAMPUNG	1	241.76	69.02	7.82	9858	96.93	70.18
9	KEP. BANGKA BELITUNG	0	88.89	70.67	7.84	12666	97.76	70.18
10	KEP. RIAU	0	260.5	74.84	9.81	13976	98.87	69.64
11	DKI JAKARTA	0	15764.3	80.47	11.05	18128	99.72	72.67
12	JAWA BARAT	0	1376.11	71.3	8.15	10790	98.48	72.66
13	JAWA TENGAH	1	1051.53	71.12	7.35	10777	93.45	74.18
14	DI YOGYAKARTA	1	1213.75	79.53	9.32	13946	94.83	74.82
15	JAWA TIMUR	1	826.38	70.77	7.39	11380	91.85	70.97
16	BANTEN	0	1313.24	71.95	8.62	11994	97.62	69.64
17	BALI	0	742.58	74.77	8.65	13886	92.98	71.68
18	NUSA TENGGARA BARAT	1	269.95	67.3	7.03	10284	87.42	65.87
19	NUSA TENGGARA TIMUR	1	110.26	64.39	7.3	7566	91.9	66.38
20	KALIMANTAN BARAT	0	33.95	66.98	7.12	8860	92.58	70.18
21	KALIMANTAN TENGAH	0	17.32	70.42	8.37	10931	99.21	69.64
22	KALIMANTAN SELATAN	0	107.96	70.17	8	12062	98.42	68.23
23	KALIMANTAN TIMUR	0	28.27	75.83	9.48	11917	98.96	73.96
24	KALIMANTAN UTARA	0	9.49	70.56	8.87	8943	95.18	72.5
25	SULAWESI UTARA	0	179.36	72.2	9.24	10731	99.87	71.26

Sebelum data dilakukan analisis dengan metode Liu, dilakukan pengecekan multikolinearitas. Berikut merupakan Tabel korelasi antarvariabel bebas dan nilai VIF.

Tabel 10. Korelasi antarvariabel bebas pada Data Kemiskinan di Indonesia

Korelasi	X1	X2	X3	X4	X5	X6
X1	1	0.5700	0.5492	0.7099	0.1661	0.2657
X2	0.5700	1	0.8215	0.8535	0.3921	0.6806
X3	0.5492	0.8215	1	0.6357	0.6445	0.3812
X4	0.7099	0.8535	0.6357	1	0.2641	0.4114
X5	0.1661	0.3921	0.6445	0.2641	1	0.2035
X6	0.2657	0.6806	0.3812	0.4114	0.2035	1

Berdasarkan Tabel diatas, dapat dilihat bahwa terdapat korelasi yang tinggi antara X4 dan X1, X2 dan X4, X2 dan X3. Hal tersebut menandakan bahwa antarvariabel memiliki hubungan yang tinggi dan memungkinkan variabel-variabel terdapat multikolinearitas.

Tabel 11. Nilai VIF pada Data Kemiskinan di Indonesia

	X1	X2	X3	X4	X5	X6
VIF	3.100	33.444	12.697	12.0287	2.4790	4.8742

Pada Tabel 11, ditunjukkan nilai VIF > 10 untuk semua variabel bebas. Hal ini menandakan bahwa data tersebut mengandung multikolinearitas.

Penaksiran Parameter β_p dengan menggunakan metode Liu

Dengan bantuan *software* R, didapatkan nilai parameter duga β_p dan SE untuk data persentase kemiskinan dengan metode Liu sebagai berikut.

Tabel 12. Nilai nilai parameter duga β_p dan SE untuk Data Persentase Kemiskinan di Indonesia

Variabel	Metode Liu		Nilai kritis uji wald
	$\hat{\beta}_p$	Nilai Wald	
$\hat{\beta}_1$	1.8014	505.2514	3.841
$\hat{\beta}_2$	-0.3443	1519.193	
$\hat{\beta}_3$	-0.3444	6471.665	
$\hat{\beta}_4$	1.8014	13390	
$\hat{\beta}_5$	-0.3443	2828.419	
$\hat{\beta}_6$	-0.3443	678.9153	

Berdasarkan Tabel 12, diperoleh nilai dugaan $\hat{\beta}_1 - \hat{\beta}_6$ metode Liu cenderung mendekati nol, sehingga penduga yang didapat lebih akurat dalam memprediksi nilai variabel terikat. Sehingga diperoleh model regresi logistik dengan metode Liu yaitu :

$$\hat{Y} = 1.8014 - 0.3443X_2 - 0.3443X_3 + 1.8014X_4 - 0.3443X_5 - 0.3443X_6$$

Selanjutnya, dilakukan pengujian parameter menggunakan uji Wald dengan nilai kritis uji Wald yaitu 3.841. Adapun hipotesis uji wald pada metode Liu sebagai berikut.

$$H_0 : \hat{\beta}_i = 0$$

$$H_1 : \hat{\beta}_i \neq 0$$

Tolak H_0 jika $W_i > X_{1,\alpha}^2$

Tidak tolak H_0 jika $W_i < X_{1,\alpha}^2$

$$X_{1,\alpha}^2 = 3.841$$

Dikarenakan nilai uji wald untuk variabel $X_1 - X_6 = W_i > X_{1,\alpha}^2$, maka dapat disimpulkan bahwa tolak H_0 . Sehingga variabel $X_1 - X_6$ yaitu kepadatan

penduduk, Indeks Pembangunan Manusia (IPM), rata-rata lama sekolah, pengeluaran per kapita, angka melek huruf, dan angka harapan hidup signifikan mempengaruhi persentase penduduk miskin di Indonesia.

2. Kesimpulan

Berdasarkan penelitian diatas, dapat disimpulkan bahwa metode Liu merupakan metode terbaik yang dapat digunakan untuk mengatasi masalah multikolinearitas, dikarenakan nilai MSE dan AIC pada metode Liu lebih kecil dibanding MSE dan AIC pada metode MLE dan LASSO. Semakin kecil nilai MSE/AIC semakin baik dalam menentukan nilai pemodelan. Dalam penelitian tersebut, variabel bebas yang digunakan berpengaruh signifikan terhadap variabel terikatnya.

DAFTAR PUSTAKA

- Montgomery, D. C., Peck, E. A., & Vining, G.G. Introduction to Linear Regression Analysis. Fourth Edition. New York: John Willey and Sons, 2012.
- Myers, R.H., 1990. Classical and Modern Regression With Application, PWSKENT publishing Company, Boston.
- N. Draper and H. Smith, Applied Regression Analysis. Jakarta. 1992.
- Fonti, V. & Belitser, E. Feature Selection using LASSO, hlm. 1-25. Research Paper In Business Analytics, VU Amsterdam. 2017
- D.W. Hosmer & S. Lemeshow. *Applied Logistics Regression*. John Wiley & Sons, New York. 2000
- M. Qasim, M. Amin, and T. Omer, "Performance Of Some New Liu Parameters for the Linear Regression Model", *Communication In Statistics Theory And Methods*. **49**(17):4178-4196.
- K. Mansson, B.M.G. Kibria, and G. Shukur, "On Liu estimators for the logit regression model", *Econ. Modell.* **29**:1483–1488. 2012.
- Liu, K. (1993) A New Class of Biased Estimate in Linear Regression. *Communications in Statistics—Theory and Methods*, **22**, 393-402.
- T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning with Sparsity the LASSO and Generalizations*. John Wiley & Sons, New York. 2015.
- R. Tibshirani, Regression Shrinkage and Selection via LASSO, *Journal of the Royal Statistical Society*. **58**(1): 267- 288
- Fonti, V. & Belitser, E. 2017. Feature Selection using LASSO, hlm. 1-25. Research Paper In Business Analytics, VU Amsterdam.
- Hastie, T., Tibshirani, R., and Wainwright, M. 2015. *Statistical Learning with Sparsity The Lasso and Generalizations*. CRC Press, New Jersey.
- Herawati, N., Nisa, K., Setiawan, E., Nusyirwan and Tiryono. 2018. Regularized Multiple Regression Methods to Deal with Severe Multicollinearity. *International Journal of Statistics and Applications*, **8**(4): 167-172.
- Hoerl, A.E. and Kennard, R.W. 2000. Ridge Regression: Biased Estimation for nonorthogonal Problems. *Technometrics*, **42**, 80-86.
- Hosmer, D.W. & Lemeshow, S. 2000. *Applied Logistics Regression*. John Wiley & Sons, New York.
- Magariños, M.G., Antoniadis, A., Cao, R., and Manteiga, W.G. 2009. Lasso Logistic Regression, Gsoft and the Cyclic Coordinate Descent Algorithm: Application to Gene Expression Data. *Statistical Applications in Genetics and Molecular Biology*. **9**(1): 8-28.
- Tibshirani, R. 1996. Regression Shrinkage and Selection via LASSO. *Journal of the Royal Statistical Society*. **58**(1): 267-288.
- Toka, O. 2016. A Comparative Study on Regression Methods in the presence of Multicollinearity. *Journal of Statisticians: Statistics and Actuarial Sciences* **2**: 47-53.

LAMPIRAN 1

BIODATA KETUA PENELITIAN

A. Data Pribadi

No	Nama Lengkap (dengan Gelar)	Prof. Ir. Netti Herawati, M.Sc., Ph.D.
	Jenis Kelamin	Perempuan
	Jabatan Fungsional	Guru Besar
	NIP	196501251990032001
	NIDN	0025016503
	Tempat dan Tanggal Lahir	Telukbetung, 25 Januari 1965
	E-mail	Netti.herawati@fmipa.unila.ac.id
	Alamat Rumah	Jl. S. Hamdani Palapa VB No 38 B. Lampung
	No Telp/HP	081273809624
	Lulusan yang telah dihasilkan	130
	Mata Kuliah yang Diampu	Statistika Dasar
		Rancangan Percobaan
		Analisis Regresi Terapan
		Metodologi Penelitian
		Nonparametrik

B. Riwayat Pendidikan

	S1	S2	S3
Nama Perguruan Tinggi	Universitas Lampung	Northern Illinois University	Gunma University
Bidang Ilmu	Ilmu Tanah	Statistika	Biometrika
Tahun masuk-lulus	1983-1987	1992-1994	1996-2000

C. . Pengalaman Penelitian dalam 5 Tahun Terakhir

No	Tahun	Judul Penelitian	Pendanaan	
			Sumber	Jmlh (jt Rph)
1	2017	Pengembangan Metode <i>Iterated Reweighted Least Trimmed Square</i> untuk Pendugaan Model <i>Generalized Estimating Equation</i> (GEE) pada Data Mengandung Pencilan	Fundamental DIKTI, ketua	50
2	2017	Pemilihan Lebar Jendela Terbaik Untuk Pendugaan Fungsi Densitas Kernel Data Waktu Bertahan (<i>Survival Time</i>) Pasien Penyakit Kanker Paru-Paru	DIPA BLU, ketua	20
3	2018	Analisis Ketegaran Regresi Robust Ridge Berdasarkan Beberapa Estimator Robust Terhadap Multikolinearitas Dan Jumlah Pencilan	DIPA BLU, Ketua	20
4	2019	Handling Full Multicollinearity And Various Numbers Of Outliers Using Robust Ridge Regression	BLU Unila, anggota	20
5	2020	Seleksi Model Logistik Biner Dengan Multikolinearitas Dan Pendugaan Peubah Yang Mempengaruhi Kemiskinan di Indonesia	DIPA BLU, ketua	20
6	2015-2020	KOMO-ISE Study	NPO International Eco-Health research Group, Anggota	200
7	2020	Gamma and Lognormal Models in the Generalized Linear Model Perspective	Hibah Penelitian dasar Unila	
8	2021	“Implementasi metode cluster nonhierarki pada pemetaan sebaran data COVID-19 di Indonesia tahun	Hibah Penelitian Terapan Unila, ketua	35
9	2021	Nonparametric Modeling Using Kernel Method for the Estimation of the Covid-9 Data in Indonesia During 2020,	Hibah BLU Unila anggota	15
10	2022			

D. Publikasi Artikel ilmiah dalam Jurnal dalam 5 Tahun Terakhir

No	Judul Artikel Ilmiah	Nama Jurnal	Vol/Nomor/Tahun
5	Handling full multicollinearity and various number of outliers using robust ridge regression.	<i>Sci. Int (Lahore), Anggota</i>	31(2):201-204/2019
6	Modelling stock return data using asymmetric volatility models; A performance comparison based on akaike information criterion and Schwarz	<i>JESR, Anggota</i>	1(1): 40-45/2019

	critierion		
7	The Effectiveness of Quantile Regression in Dealing with Potential Outliers	<i>BAREKENG, Jurnal Ilmu Matematika dan Terapan, Ketua</i>	14(2):305-312/2020
8	Gamma and Lognormal Models in the Generalized Linear Model Perspective	<i>IJIRSET, anggota</i>	9(9): 8478-8488, 2020
9	Selecting the Method to Overcome Partial and Full Multicollinearity in Binary Logistic Model	<i>International Journal of Statistics and Applications. Ketua</i>	10(3): 55-59, 2020
10	Smoking has physical and psychoactive effects, and heavy smoking is associated with depression	<i>Japan Journal of health & Human Ecology, ketua</i>	86(6): 276-280, 2020
11	Robust Principal Component Trimmed Clustering of Indonesian Provinces Based on Human Development Index Indicators	<i>Journal of Phys.: Conf. Ser. 1751 012021, anggota</i>	2021
12	Structural Equation Modeling of Factors Affecting Student Motivation in Thesis Preparation	<i>Journal of Phys.: Conf. Ser. 1751 012021, anggota</i>	2021
13	Modeling with generalized linear model on covid-19: Cases in Indonesia	<i>International Journal of Electronics and Communications System. Volume 1, Issue 1, 25-32. anggota</i>	Volume 1, Issue 1, 25-32. 2021
14	Nonparametric Modeling Using Kernel Method for the Estimation of the Covid-9 Data in Indonesia During 2020	<i>International Journal of Mathematics Trends and Technology ,</i>	Volume 67 Issue 6, 136-144, June, 2021
15	Smooth Version of MWW statistics, (2022). <i>AXIOMS</i> . Vol. 11 No.7 2022 pp 1-9	<i>AXIOMS</i>	. Vol. 11 No.7 2022 pp 1-9
16	Convergence Rates of Maxima of Non-identical Sums , (2022).	<i>Pak.j.stat.oper.res.</i>	Vol. 18 No.1 2022 pp 245-248
17	Overdispersion Data Modeling Cases of Filariasis in East Lampung Province, Indonesia.	<i>International Journal of Innovative Science and Research Technology</i>	Volume 8, Issue 6, June 2023
18	A High-Order Fuzzy Time Series Forecasting Method Based on an Intersection Operation for Forecasting USD Foreign Exchange	<i>International Journal of Latest Engineering Research and Applications (IJLERA).</i>	Volume – 08, Issue – 01, January 2023, PP – 46-53
19	Environmental-based Diseases of students in urban and rural areas, Lampung province , Indonesia	<i>International journal of innovative research and scientific</i>	Volume 6 no 2 pp 242-249, June 2023
20	Poisson Ridge Regression for Multicollinearity Data: Case Study of the Number of Maternal Deaths in Lampung Province, Indonesia.	<i>International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET).</i>	Volume 13, Issue 1, January 2024

BIODATA ANGGOTA PENELITI 1

A. Identitas Diri

1	Nama Lengkap (dengan Gelar)	Dr. Khoirin Nisa, M.Si
2	Jenis Kelamin	Perempuan
3	Jabatan Fungsional	Lektor
4	NIP	197407262000032001
5	NIDN	0026077401
6	Tempat dan Tanggal Lahir	Jakarta, 26 Juli 1974
7	E-mail	Khoirin.nisa@fmipa.unila.ac.id
8	Alamat Rumah	Perumahan Griya Kencana Blok C No. 6 Rajabasa Bandar-Lampung
9	Nomor Telepon/HP	081379846402
10	Alamat Kantor	Jurusan Matematika FMIPA Unila
11	Nomor Telepon/Fax	(0721)-701609
12	Lulus yang Telah Dihasilkan	S160 orang
13	Mata Kuliah yang Diampu	1. Analisis Multivariat 2. Statistika Matematika 3. Matriks untuk Statistika 4. Statistika Dasar

B. Riwayat Pendidikan

	S1	S2	S3
Nama Perguruan Tinggi	Unpad Bandung	IPB	IPB – Universitas Franche Comte (<i>double degree</i>)
Bidang Ilmu	Matematika	Statistika	Statistika – <i>Mathematics</i>
Tahun Masuk-Lulus	1992-1997	1998-2001	2011-2015
Judul Skripsi/ Thesis/ Disertasi	Masalah Nilai Batas Persamaan Gelombang Satu Dimensi Pada	Analisis Kausal dengan Peubah Laten untuk Data Kategorik	<i>On Multivariate Dispersion Analysis</i>
Nama Pembimbing/ Promotor	Drs. Wiratmadja, M.S.	Prof. Asep Saefuddin, Dr. Mulya Siregar	Prof. Asep Saefuddin, Prof. I Wayan Mangku, Dr. Aji H. Wigena, Prof. Celestin C. Kokonendji

C. Publikasi Artikel Ilmiah dalam Jurnal dalam 5 Tahun Terakhir

No	Judul Artikel Ilmiah	Nama Jurnal / Proceeding	Volume/ Nomor/Tahun
1	Characterization of multivariate normal Poisson models	<i>Journal of Iranian Statistical Society</i>	14/2/2015
2	Generalized variance estimation of normal Poisson models	<i>Springer Proceeding of Mathematics and Statistics</i>	124/ 21/ 16
3	Empirical comparison of the generalized variance estimators of some normal stable Tweedie models: a simulation study	<i>Applied Mathematical Sciences</i>	10/63/2016
4	Robust generalized estimating equation when data contain outliers	<i>INSIST</i>	02/01/2017
5	On generalized variance of normal Poisson model and Poisson variance estimation under Gaussianity	<i>ARPJ Journal of Engineering and Applied Sciences</i>	12/12/2017
6	A robust procedure for GEE model	<i>Far East Journal of Mathematical Sciences</i>	102/03/2017
4	<i>The Optimal Bandwidth for Kernel Density Estimation of Skewed Distribution: A Case Study on Survival Data of Cancer Patients</i>	Prosiding seminar nasional metode kuantitatif. Issn 978-602-98559-3-7	24 – 25/11/2017

D. Pemakalah Seminar Ilmiah (oral presentation) dalam 5 Tahun Terakhir

No	Nama Pertemuan Ilmiah/ Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	<i>International Conference on Computational Mathematics, Computational Geometrics and Statistics.</i>	<i>Generalized Variance Estimation of Normal Poisson Models</i>	3-4 Februari 2014, Singapura
2	<i>International Conference of Sciences Technology and Interdisciplinary Research</i>	<i>Robust generalized estimating equation when data contain outliers</i>	23-24 Agustus 2016, Bandar Lampung

BIOADATA ANGGOTA PENELITI 2

A. Data Pribadi

No	Nama Lengkap (dengan Gelar)	Subian Saidi, S.Si., M.Si.
	Jenis Kelamin	Laki-laki
	Jabatan Fungsional	Asisten Ahli
	NIP	19581101 198803 1 002
	NIDN	0001116901
	E-mail	erstat@gmail.com
	Alamat Rumah	Jl. Asoka B-60 Perumahan Bataranila Hajimena, Natar, Lampung Selatan, 35144
	No Telpon/HP	08127270312
	Lulusan yang telah dihasilkan	130
	Mata Kuliah yang Diampu	Statistika Dasar
		Metode Statistika
		Teknik Sampling
		Kapita Selekt Statistika
		Nonparametrik

B. Riwayat Pendidikan

	S1	S2
Nama Perguruan Tinggi	Unievrstas Lampung	Universitas Indonesia
Bidang Ilmu	Statistika	Matematika
Tahun masuk-lulus	1981-1986	2000-2003

C.

Pengalaman Penelitian

No	Tahun	Judul Penelitian	Sumber Dana	Status
1	2006	<i>Robust principal Component Analysis using Minimum Covariance Determinant Estimator</i>	PHK A2	Anggota
2	2007	Indeks kecocokan dari beberapa metode estimasi untuk ukuran sampel tertentu pada model persamaan struktural	PHK A2	Ketua
3	2008	Pengaruh langsung dan tak langsung dari peubah laten pada model persamaan struktural	DIPA PNBP UNILA	Ketua
4	2016	Model persamaan struktural pada data ordinal	Mandiri	Ketua
5	2017	Analisis factor konfirmatori dengan bantuan SEM-PLS untuk analisis data ordinal (setudi kasus penilaian pejabat Unila tahun 2016)	Mandiri	Ketua

D. Publikasi Artikel ilmiah dalam Jurnal dalam 5 Tahun Terakhir

No	Judul Artikel Ilmiah	Nama Jurnal	Vol/Nomor/Tahun
1	Model persamaan struktural untuk analisis data (studi kasus survey kepuasan konsumen)	<i>Prosiding Semirata 2016 Bidang MIP BKS PTN wil barat, Unsri. ISBN: 978-602-71798-1-3</i>	22-24 Mei 2016
2	<i>Regularized Multiple Regression Methods to Deal with Severe Multicollinearity</i>	International Journal of Statistics and Applications	8(4): 167-172/2018
3	<i>Analisis Regresi Nonparametrik Dengan Teknik Smoothing</i>	Semirata, Medan, 5 Mei 2018	2018
4	<i>Metode Estimasi Diagonal Weighted Least Square (DWLS) Untuk Berbagai Ukuran Sampel (Studi Kasus Kualitas Pelayanan Perpustakaan Unila)</i>	Prosiding seminar nasional metode kuantitatif II. Issn 978-623-90150-0-8	19-20 November 2018
5	<i>Handling Full Multicollinearity And Various Numbers Of Outliers Using</i>	Sci.Int.(Lahore)	31(2),201-204/2019

	<i>Robust Ridge Regression</i>		
6	<i>Modeling Stock Return Data using Asymmetric Volatility Models : A Performance Comparison based on the Akaike Information Criterion and Schwarz Criterion</i>	Journal of Engineering and Scientific Research (JESR), pISSN: 268-0338; eISSN: 268-1695	1(1) 40-45/2019

E. Pemakalah Seminar Ilmiah (*oral presentation*) dalam 5 Tahun Terakhir

No	Nama Pertemuan Ilmiah/Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1	SEMIRATA	SEM untuk Studi kasus Survey Kepuasan Konsumen	<i>UNSRI, 22-24 Mei 2016</i>
2	<i>Analisis Regresi Nonparametrik Dengan Teknik Smoothing</i>	Semirata, Medan, 5 Mei 2018	2018
3	seminar nasional metode kuantitatif.	<i>Metode Estimasi Diagonal Weighted Least Square (DWLS) Untuk Berbagai Ukuran Sampel (Studi Kasus Kualitas Pelayanan Perpustakaan Unila)</i>	<i>Bandar Lampung, 2018</i>
4	seminar nasional metode kuantitatif.	<i>Metode Estimasi Diagonal Weighted Least Square (DWLS) Untuk Berbagai Ukuran Sampel (Studi Kasus Kualitas Pelayanan Perpustakaan Unila)</i>	<i>Bandar Lampung, 2019</i>