

**MODEL REGRESI MENGGUNAKAN *LEAST ABSOLUTE
SHRINKAGE AND SELECTION OPERATOR (LASSO)* PADA
DATA BANYAKNYA PENDERITA GIZI BURUK
KABUPATEN/KOTA DI JAWA TENGAH**



SKRIPSI

Disusun Oleh: AULIA

PUTRI ANDANA

24010212140067

**DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2016

**MODEL REGRESI MENGGUNAKAN *LEAST ABSOLUTE
SHRINKAGE AND SELECTION OPERATOR (LASSO)* PADA
DATA BANYAKNYA PENDERITA GIZI BURUK
KABUPATEN/KOTA DI JAWA TENGAH**

Disusun Oleh:

Aulia Putri Andana

24010212140067

Skripsi

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Sains pada
Departemen Statistika Fakultas Sains dan Matematika Universitas Diponegoro**

**DEPARTEMEN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2016

HALAMAN PENGESAHAN I

Judul : Model Regresi Menggunakan *Least Absolute Shrinkage and Selection Operator* (LASSO) pada Data Banyaknya Penderita Gizi Buruk Kabupaten/Kota di Jawa Tengah

Nama : Aulia Putri Andana

NIM : 24010212140067

Telah diujikan pada ujian Tugas Akhir dan dinyatakan lulus pada tanggal 24 Oktober 2016.

Semarang, Oktober 2016

Mengetahui,

Ketua Departemen Statistika
Fakultas Sains dan Matematika



Dra. Dwi Ispryanti, M.Si.
NIP. 195709141986032001

Panitia Penguji Ujian Tugas Akhir
Ketua

Dra. Suparti, M.Si.
NIP. 196509131990032001

HALAMAN PENGESAHAN II

Judul : Model Regresi Menggunakan *Least Absolute Shrinkage and Selection Operator* (LASSO) pada Data Banyaknya Penderita Gizi Buruk Kabupaten/Kota di Jawa Tengah

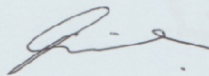
Nama : Aulia Putri Andana

NIM : 24010212140067

Telah diujikan pada ujian Tugas Akhir pada tanggal 24 Oktober 2016

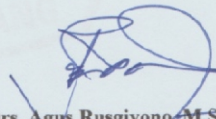
Semarang, Oktober 2016

Dosen Pembimbing I



Diah Safitri, S.Si., M.Si.
NIP. 197510082003122001

Dosen Pembimbing II



Drs. Agus Rusgivono, M.Si.
NIP. 196408131990011001

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT yang telah melimpahkan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan tugas akhir yang berjudul **“Model Regresi Menggunakan *Least Absolute Shrinkage and Selection Operator* (LASSO) pada Banyaknya Penderita Gizi Buruk Kabupaten/Kota di Jawa Tengah”**. Tugas akhir ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Strata Satu (S1) pada Departemen Statistika Fakultas Sains dan Matematika Universitas Diponegoro.

Pada kesempatan ini penulis ingin menyampaikan terima kasih kepada:

1. Drs. Tarno, M.Si. selaku Ketua Departemen Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
2. Diah Safitri, S.Si, M.Si selaku dosen pembimbing I dan Drs. Agus Rusgiyono, M.Si. selaku dosen pembimbing II
3. Semua Bapak dan Ibu Dosen Departemen Statistika FSM Universitas Diponegoro
4. Semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu hingga terselesaikannya penulisan tugas akhir ini.

Kritik dan saran dari pembaca akan menjadi masukan yang sangat berharga. Harapan penulis semoga tugas akhir ini dapat bermanfaat bagi penulis khususnya dan bagi pembaca pada umumnya.

Semarang,

Penulis

ABSTRAK

Gizi buruk adalah bentuk terparah dari proses terjadinya kekurangan gizi yang menahun. Gizi buruk dipengaruhi oleh banyak faktor yang saling terkait. Dalam penelitian ini, dilakukan pemodelan dari faktor-faktor yang mempengaruhi gizi buruk menggunakan metode *Least Absolute Shrinkage Selection and Operator* (LASSO) dengan algoritma *Least Angle Regression* (LARS) karena pada faktor-faktor yang mempengaruhi gizi buruk terdeteksi multikolinearitas. LASSO menyusutkan koefisien regresi dari variabel bebas yang memiliki korelasi tinggi menjadi tepat pada nol atau mendekati nol. Koefisien LASSO dicari dengan menggunakan pemrograman kuadratik sehingga digunakan algoritma LARS yang lebih efisien dalam komputasi LASSO. Berdasarkan analisis yang telah dilakukan, model LASSO pada data gizi buruk Kabupaten/Kota di Jawa Tengah tahun 2014 diperoleh pada tahap kedua saat nilai $s=0,02$ dengan nilai MSE sebesar 0,82977. Disimpulkan bahwa variabel bayi (0-6 Bulan) yang diberi ASI Eksklusif, rumah tangga berperilaku hidup bersih dan sehat, bayi yang mendapat imunisasi Hepatitis B, bayi yang mendapat imunisasi DPT-HB3, rumah dengan sanitasi yang layak, dan rumah dengan air minum sesuai dengan syarat kesehatan berpengaruh terhadap bayi gizi buruk di Jawa Tengah tahun 2014.

Kata Kunci: *gizi buruk, multikolinearitas, LASSO, LARS*

ABSTRACT

Malnutrition is the most severe form of the occurrence of chronic malnutrition. Malnutrition is influenced by many interrelated factors. In this study, carried out the modeling of the factors that influence malnutrition using Least Absolute Shrinkage and Selection Operator (LASSO) method with Least Angle Regression (LARS) algorithms due to the factors that influence malnutrition there is multicollinearity detected. LASSO shrinks the regression coefficients of the independent variables that have a high correlation to be right at zero or close to it. LASSO coefficients calculated using quadratic programming so that LARS algorithm is used due to efficiency on LASSO computing. Based on the analysis performed, the model of LASSO in malnutrition data at Central Java Regency/City in 2014 was obtained in the second stage when the value $s = 0,02$ with MSE value of 0,82977. Concluded that the infants variable (0-6 months) which got exclusive breastfeeding, household that behave with clean and healthy life, infants immunized against Hepatitis B, baby immunized against DPT-HB3, house with proper sanitation, and house with drinking water which accordance with health requirements affect the infant malnutrition in Central Java in 2014.

Keywords: *malnutrition, multicollinearity, LASSO, LARS*

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN I.....	ii
HALAMAN PENGESAHAN II.....	iii
KATA PENGANTAR	iv
ABSTRAK	v
ABSTRACT.....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	x
DAFTAR LAMPIRAN.....	xi
BAB I. PENDAHULUAN	
1.1 Latar Belakang.....	1
1.3 Rumusan Masalah.....	2
1.3 Pembatasan Masalah.....	3
1.4 Tujuan Penelitian	3
BAB II TINJAUAN PUSTAKA	
2.1 Gizi buruk	4
2.2 Penyebab Gizi Buruk	5
2.2.1 Persediaan Pangan	5
2.2.2 Pola Asuh Anak	5
2.2.3 Pelayanan Kesehatan Dasar	6
2.2.4 Sanitasi dan Air Bersih	7

2.3 Analisis Regresi Berganda	8
2.4 Metode Kuadrat Terkecil	9
2.5 Multikolinearitas	10
2.6 <i>Least Absolute Shrinkage and Selection Operator</i> (LASSO).....	12
2.7 <i>Least Angle Regression</i> (LARS).....	13
2.7.1 Langkah-langkah estimasi koefisien LASSO dengan algoritma LARS	14
2.7.2 Modifikasi dari <i>Least Angle Regression</i>	15
2.8 Validasi Silang	16
 BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data.....	18
3.2 Variabel Penelitian.....	18
3.3 Langkah-langkah Analisis	19
3.4 Diagram Alir (<i>Flow Chart</i>) Penelitian.....	20
 BAB IV HASIL DAN PEMBAHASAN	
4.1 Deskripsi Data.....	21
4.2 Analisis Regresi Linear Berganda	22
4.3 Uji Multikolinearitas	22
4.4 <i>Least Absolute Shrinkage and Selection Operator</i> (LASSO) dengan algoritma LARS	23
BAB V KESIMPULAN	39
DAFTAR PUSTAKA	40
LAMPIRAN.....	42

DAFTAR GAMBAR

	Halaman
Gambar 1. Diagram Alur Analisis	20
Gambar 2. Plot tahapan seleksi variabel untuk menduga koefisien LASSO dengan algoritma LARS	29
Gambar 3. Nilai validasi silang dengan menggunakan mode <i>fraction</i>	32

DAFTAR TABEL

	Halaman
Tabel 1. Klasifikasi Status Gizi Anak	5
Tabel 2. Variabel-variabel Penelitian.....	23
Tabel 3. Statistika Deskriptif dari Variabel Penelitian.....	24
Tabel 4. Nilai VIF dari Setiap Variabel Prediktor	25
Tabel 5. Hasil Calon Koefisien LASSO dengan Algoritma LARS.....	37
Tabel 6. Hasil Nilai s dan Nilai CV MSE.....	38

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Nilai SD badan menurut umur (BB/U).....	43
Lampiran 2. Data Penelitian.....	47
Lampiran 3. Variabel yang telah distandarkan dengan rata-rata nol dan ragam satu.....	49
Lampiran 4. Output Regresi Linear Berganda Menggunakan OLS.....	51
Lampiran 5. Nilai Koefisien Setiap Tahapan.....	52
Lampiran 6. Hasil Perhitungan Tahap Kedua dimana X_1 sebagai Variabel yang Terseleksi.....	53
Lampiran 7. Hasil Perhitungan Tahap Kedua dimana X_9 sebagai Variabel yang Terseleksi.....	55
Lampiran 8. Hasil Perhitungan Tahap Ketiga dimana X_8 sebagai Variabel yang Terseleksi.....	57
Lampiran 9. Hasil Perhitungan Tahap Keempat dimana X_2 sebagai Variabel yang Terseleksi.....	59
Lampiran 10. Hasil Perhitungan Tahap Kelima dimana X_3 sebagai Variabel yang Terseleksi.....	62
Lampiran 11. Hasil Perhitungan Tahap Keenam dimana X_7 sebagai Variabel yang Terseleksi.....	65
Lampiran 12. Hasil Perhitungan Tahap Ketujuh dimana X_5 sebagai Variabel yang Terseleksi.....	68

Lampiran 13. Hasil Perhitungan Tahap Kedelapan dimana X_6 sebagai Variabel yang Terseleksi.....	71
Lampiran 14. Hasil Perhitungan Tahap Kesembilan dimana X_4 sebagai Variabel yang Terseleksi.....	74
Lampiran 15. Nilai s atau $\frac{\sum \hat{\beta}_j }{\max\sum \hat{\beta}_j }$ untuk Setiap Tahapan.....	77
Lampiran 16. Hasil Pengolahan Data.....	78

BAB I

PENDAHULUAN

1.1 Latar Belakang

Gizi buruk adalah bentuk terparah dari proses terjadinya kekurangan gizi yang menahun (Depkes RI, 2005). Persoalan gizi menjadi salah satu indikator penting yang menjadi kesepakatan global dalam *Milleneum Development Goals* (MDGs). Setiap negara secara bertahap harus mampu mengurangi jumlah balita yang bergizi buruk atau gizi kurang sehingga mencapai 15 persen pada tahun (Saputra dan Nurrizka, 2012).

Provinsi Jawa Tengah memiliki jumlah gizi buruk dengan indikator berat badan menurut tinggi badan sebanyak 3,942 balita atau 0,16% persen dari jumlah balita yang ada di Jawa Tengah pada tahun 2014, angka ini masih lebih rendah dari target nasional sebesar 3% (Dinkes, 2014). Meskipun jumlah gizi buruk masih berada di bawah target nasional, akan lebih baik jika di Jawa Tengah tidak terdapat penderita jumlah gizi buruk.

Menurut *UNICEF* (1998), gizi buruk dipengaruhi oleh beberapa penyebab tidak langsung antara lain persediaan pangan, pola asuh anak, pelayanan kesehatan dasar, serta sanitasi dan air bersih. Gizi buruk dipengaruhi oleh banyak faktor yang saling terkait (Depkes RI, 2005).

Pada penulisan tugas akhir ini dilakukan pemodelan regresi dari faktor-faktor penyebab banyaknya penderita gizi buruk di Jawa Tengah berdasarkan Kabupaten/Kota tahun 2014. Adapun variabel respon yang digunakan adalah banyaknya penderita gizi buruk dan variabel prediktor yang digunakan adalah pemberian ASI Eksklusif, rumah tangga berperilaku hidup bersih dan sehat,

imunisasi dasar yang meliputi imunisasi BCG, imunisasi Polio, imunisasi Campak, imunisasi Hepatitis B, dan imunisasi DPT-HB3, sanitasi layak, dan air minum yang memenuhi syarat kesehatan.

Gizi buruk dipengaruhi oleh banyak faktor yang saling terkait sehingga pada pemodelan regresi faktor-faktor yang mempengaruhi banyaknya penderita gizi buruk rentan terdeteksi masalah multikolinearitas. Menurut Gujarati dan Porter (2009), multikolinearitas dapat menyebabkan kesulitan dalam mendapatkan estimasi yang tepat.

Menurut Tibshirani (1996), metode penyusutan *Least Absolute Shrinkage and Selection Operator* (LASSO) dapat digunakan dalam memodelkan regresi yang mengandung multikolinearitas karena LASSO memiliki model regresi yang lebih mudah untuk diinterpretasikan. Menurut Hastie *et al.* (2008), algoritma *Least Angle Regression* (LARS) dapat digunakan untuk mempermudah komputasi LASSO karena LARS mempunyai modifikasi algoritma yang lebih efisien.

Dalam penulisan tugas akhir ini model regresi dibentuk dari faktor-faktor yang mempengaruhi banyaknya penderita gizi buruk di Jawa Tengah berdasarkan Kabupaten/Kota tahun 2014 yang terdeteksi masalah multikolinearitas menggunakan metode pendekatan *Least Absolute Shrinkage and Selection Operator* (LASSO) dengan algoritma *Least Angle Regression* (LARS).

1.2 Rumusan Masalah

Bagaimana membentuk model regresi menggunakan metode LASSO dengan algoritma LARS terhadap faktor-faktor yang mempengaruhi gizi buruk Kabupaten/Kota di Jawa Tengah tahun 2014 yang terdapat masalah

multikolinearitas.

1.3 Pembatasan Masalah

Pada penelitian ini, masalah yang dibatasi adalah sebagai berikut:.

1. Data yang digunakan adalah data gizi buruk dari setiap Kabupaten/Kota yang berada di Jawa Tengah dengan jumlah 35 Kabupaten/Kota pada tahun 2014
2. Dalam pemodelan faktor-faktor yang mempengaruhi data gizi buruk Kabupaten/Kota di Jawa Tengah tahun 2014 metode yang digunakan adalah metode LASSO dengan algoritma LARS.

1.4 Tujuan Penelitian

Tujuan dari penelitian dalam tugas akhir ini adalah membentuk model regresi menggunakan metode LASSO dengan algoritma LARS terhadap faktor-faktor yang mempengaruhi gizi buruk Kabupaten/Kota di Jawa Tengah tahun 2014 yang terdapat masalah multikolinearitas.