

Received: 18 May 2022

Revised: 27 June 2022

Accepted: 28 June 2022

Published: 30 June 2022

Penerapan Regresi *Least Absolute Shrinkage and Selection Operator* (LASSO) Untuk Mengidentifikasi Variabel yang Berpengaruh terhadap Kejadian *Stunting* di Indonesia

Tesa Trilonika Pardede ^{1, a)}, Bagus Sumargo ^{2, b)}, Widyanti Rahayu ^{3, c)}

¹*Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. Jl. Rawamangun Muka, Kota Jakarta Timur, DKI Jakarta, 13220.*

Email:^{a)} tesatrilonika00@gmail.com, ^{b)} bagussumargo@unj.ac.id, ^{c)} widyanti.rahayu@gmail.com

Abstract

Linear regression analysis is an analytical method that works to analyze data and draw meaningful conclusions about one variable against other variables. Some assumptions that must be met in linear regression analysis are normal distribution, there is no correlation between errors. Constraints that cause assumptions are not met, for example the occurrence between independent variables (multicollinearity). The analysis in this study uses the Least Absolute Shrinkage and Selection Operator (LASSO) regression method with the Least Angle Regression (LAR) algorithm because the stunting data in Indonesia has multicollinearity problems among the independent variables used. LASSO solves the case of multicollinearity while at the same time it is possible to shrink the regression coefficient from the highly correlated independent variable to exactly zero. The LASSO coefficient obtained uses quadratic so that the LAR algorithm is used which is more efficient in LASSO computing. The conclusion obtained based on the analysis that has been carried out, namely exclusive breastfeeding (X1), protein consumption (X2), DPT-HB exercise (X5), maternal height (X8) and diarrhea (X9) have an effect on stunting in Indonesia in 2018.

Keywords: stunting, multicollinearity, LASSO, LAR.

Abstrak

Analisis regresi linier merupakan metode analisis yang berfungsi untuk menganalisis data serta menarik kesimpulan yang berarti tentang ketergantungan suatu variabel terhadap variabel yang lain. Beberapa asumsi yang harus dipenuhi dalam analisis regresi linier yaitu distribusi normal, tidak ada korelasi antara galat. Kendala yang menyebabkan asumsi tidak terpenuhi, misalkan terjadinya korelasi diantara variabel bebas (multikolinieritas). Analisis pada penelitian ini menggunakan metode regresi *Least Absolute Shrinkage And Selection Operator* (LASSO)

dengan algoritma *Least Angle Regression* (LAR) dikarenakan pada data *stunting* di Indonesia terdapat adanya masalah multikolinieritas di antara variabel bebas yang digunakan. LASSO menyelesaikan kasus multikolinieritas pada regresi sekaligus memungkinkan untuk menyusut koefisien regresi dari variabel bebas yang berkorelasi tinggi sampai tepat nol. Koefisien LASSO didapatkan menggunakan pemrograman kuadratik sehingga digunakan algoritma LAR yang lebih efisien dalam komputasi LASSO. Kesimpulan yang diperoleh berdasarkan analisis yang telah dilakukan yaitu variabel asi eksklusif (X1), konsumsi protein (X2), imunisasi DPT-HB (X5), tinggi badan ibu (X8) dan diare (X9) berpengaruh terhadap *stunting* di Indonesia tahun 2018.

Kata-kata kunci: *stunting*, multikolinieritas, LASSO, LAR

PENDAHULUAN

Stunting adalah masalah kekurangan gizi kronis yang dikarenakan kurangnya asupan makanan yang berkepanjangan dan berdampak pada pertumbuhan anak, yaitu tinggi badan anak lebih rendah atau lebih pendek dari usia standar (Riskesdas, 2018). Masalah gizi ini lebih beresiko jika sudah terjadi sejak di dalam kandungan. *Stunting* dipengaruhi oleh faktor-faktor yang berhubungan langsung dengan tumbuh kembang bayi atau balita, seperti pola asuh, pemberian ASI eksklusif, pemberian MP-ASI, pemberian imunisasi lengkap, kecukupan protein dan mineral, penyakit infeksi, dan genetik (Aridiyah et al, 2015). Faktor-faktor yang signifikan mempengaruhi *stunting* diantaranya tingkat asupan energi, berat badan lahir, tingkat pendidikan orang tua, dan tingkat pendapatan (Setiawan et al, 2018). Menurut UNICEF (1990) faktor lainnya adalah pengetahuan ibu yang kurang, pola asuh yang salah, sanitasi atau kebersihan yang buruk dan rendahnya pelayanan kesehatan. Secara tidak langsung penyebab *stunting* secara eksternal dipengaruhi oleh faktor sosial ekonomi keluarga, seperti tingkat pendidikan ibu, status pekerjaan ibu, dan pendapatan keluarga (Aridiyah et al, 2015).

Riset Kesehatan Dasar (Riskesdas) adalah riset secara nasional yang dilakukan setiap lima tahun sekali oleh Badan Litbangkes Kementerian Kesehatan Republik Indonesia. Riskesdas dilakukan untuk menilai perkembangan status kesehatan masyarakat, faktor risiko, dan perkembangan upaya pembangunan kesehatan. Menurut Riset Kesehatan dasar (2018) Indonesia memiliki 30,8% atau 7,3 juta anak di Indonesia mengalami *stunting* yang artinya satu dari tiga balita mengalami *stunting*. Menurut Riskesdas (2018), angka *stunting* di Indonesia meningkat dari 27,5% (2016) menjadi 29,6% (2017) dan pada tahun 2018 mencapai angka 30,8%, lebih tinggi dari target Badan Kesehatan Dunia (WHO) yang menetapkan angka *stunting* seharusnya tidak lebih dari 20%.

Menurut Picauly & Toy (2013) dampak yang diakibatkan *stunting* yaitu kognitif lemah, perkembangan fisik dan psikomotorik terhambat, penurunan prestasi akademik dan peningkatan risiko penyakit degeneratif. Menurut Hoffman, et al (2000) dampak yang diakibatkan *stunting* akan meningkatkan risiko obesitas dan akan lebih rentan terhadap penyakit tidak menular. Penelitian lebih lanjut telah dilakukan untuk mengetahui faktor-faktor yang mempengaruhi *stunting*. Penelitian Setiawan, et al (2018) yang membahas mengenai faktor-faktor yang berhubungan dengan kejadian *Stunting* di wilayah kerja puskesmas andalas Kecamatan Padang Timur Kota Padang Tahun 2018 menggunakan 10 variabel bebas yang berkaitan dengan *stunting* dengan metode uji Chi-square dan uji regresi logistik berganda.

Regresi linier berganda adalah suatu metode yang dapat digunakan untuk menganalisis faktor-faktor yang mempengaruhi suatu variabel terikat dan untuk meramalkan pengaruh

dua variabel bebas atau lebih terhadap satu variabel terikat. Ketika dihadapkan dengan kondisi dimana data yang ada terjadi kasus multikolinieritas, maka regresi linier berganda dikatakan tidak cukup mampu untuk memodelkan dan menganalisis data tersebut. *Stunting* dipengaruhi oleh banyak faktor yang saling berkaitan sehingga pada pemodelan regresi faktor-faktor yang mempengaruhi *stunting* rentan terdeteksi masalah multikolinieritas. Multikolinieritas adalah masalah yang sering terjadi dalam analisis regresi berganda di mana terdapat hubungan yang signifikan di antara dua atau lebih variabel bebas. Multikolinieritas mengakibatkan adanya penduga kuadrat terkecil yang memiliki variansi besar (Montgomery & Peck, 1992). Multikolinieritas akan mempengaruhi keakuratan prediksi model dan menyebabkan kesalahan dalam pengambilan keputusan. Akurasi prediksi terkadang dapat ditingkatkan dengan mengecilkan nilai koefisien regresi menjadi nol (Hastie et al, 2008). Beberapa metode yang dapat digunakan untuk mengatasi masalah multikolinieritas antara lain metode regresi ridge (gulud) dan *least absolute shrinkage and selection operator* (LASSO).

LASSO merupakan teknik untuk menyeleksi variabel pada data yang berdimensi besar dan dapat menyusutkan koefisien regresi menjadi tepat nol (Tibshirani, 1996). Perbedaan antara regresi LASSO dan regresi ridge adalah dalam mengestimasi koefisien regresi, dimana koefisien regresi ridge hanya dapat disusutkan mendekati nol, sedangkan koefisien regresi LASSO memungkinkan untuk menyusut sampai tepat nol. Kelebihan dari regresi LASSO, yaitu dapat digunakan untuk menyeleksi variabel bebas dalam model, sehingga hanya variabel berpengaruh yang dimasukkan dalam model dan memudahkan dalam menginterpretasikan model regresi. Penalti yang paling terkenal adalah metode LASSO karena perhitungannya cepat untuk menyelesaikan masalah pengoptimalan kecembungan (Santi et al, 2019). Perhitungan yang tepat dalam metode LASSO untuk menjadi lebih baik adalah menggunakan algoritma LAR. Least Angle Regression (LAR) merupakan metode yang dimana algoritmanya dapat diubah ke dalam suatu komputasi algoritma pada metode LASSO. Modifikasi algoritma LAR untuk metode LASSO mengarah kepada efisiensi dalam memperkirakan koefisien LASSO.

Penelitian mengenai LASSO sebelumnya pernah dilakukan oleh Santi, et al (2019) dengan judul *Variable Selection Methods for Generalized Linear Mixed Models by Penalized Lasso in Modeling the Scores of Indonesian Students in Programme for International Student Assessment (PISA)*. Penelitian yang dilakukan oleh Dewi (2010) dengan judul *Metode OLS, LASSO, dan PLS Pada Data yang Mengandung Adanya Multikolinieritas* dan penelitian Prabowo, et al (2015) tentang *Pemodelan Pertumbuhan Ekonomi Jawa Tengah Menggunakan Metode Pendekatan Least Absolute Shrinkage and Selection Operator (LASSO)*.

METODOLOGI

Bahan dan Data

Pada penelitian ini, data yang digunakan adalah data sekunder yang diperoleh dari website resmi Badan Pusat Statistik (BPS) dan riset kesehatan dasar (Riskesdas) yang dilaksanakan oleh Badan Litbangkes Kemenkes RI. Data yang digunakan adalah data yang berkaitan dengan *stunting* yang meliputi unit amatan sebanyak 34 Provinsi di Indonesia pada tahun 2018. Variabel terikat dalam penelitian ini adalah *stunting* di Indonesia. Selain itu, variabel bebas yang digunakan dalam penelitian ini adalah Asi eksklusif (X1), Konsumsi protein (X2), Imunisasi Hepatitis B (X3), Imunisasi *Bacillus Calmette-Guerin*

(X4), Imunisasi DPT-HB (X5), Imunisasi Polio (X6), Imunisasi Campak (X7), Tinggi badan ibu (X8), Diare (X9), ISPA (X10), Sumber air minum layak (X11).

Metode Penelitian

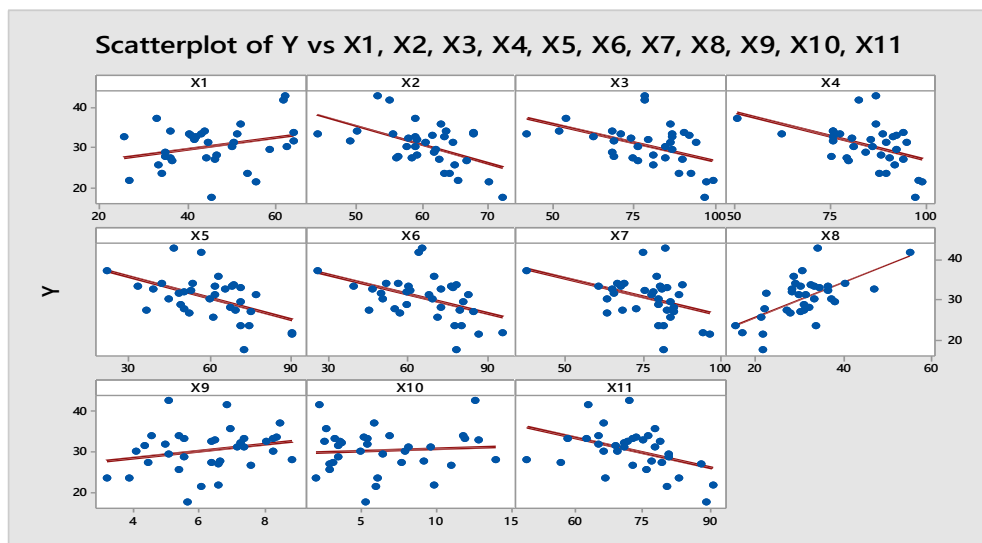
Penelitian ini dilakukan melalui penelitian pustaka, yaitu melalui buku teks pendukung penelitian dan karya tulis ilmiah yang diterbitkan di jurnal. Guna mempermudah penghitungan dan memperoleh hasil yang akurat, penulis menggunakan software RStudio. Tahapan yang akan dilakukan adalah sebagai berikut:

1. Melakukan eksplorasi data berupa plot di antara variabel bebas dan variabel terikat, korelasi antar variabel, serta melakukan statistika deskriptif.
2. Melakukan analisis regresi berganda menggunakan metode kuadrat terkecil (MKT).
3. Mendeteksi adanya data yang mengandung multikolinieritas dengan menggunakan Variance Inflation Factor (VIF) dan melakukan uji asumsi klasik dalam regresi.
4. Melakukan analisis regresi dengan metode LASSO.
 - a) Menentukan pendugaan untuk koefisien regresi LASSO pada setiap tahapan yang dilakukan .
 - b) Mencari dan memilih nilai s optimal menggunakan proses validasi silang (CV).
 - c) Menduga koefisien model terbaik dengan cara melihat nilai CV MSE minimum yang diperoleh dari proses validasi silang.

HASIL DAN PEMBAHASAN

Eksplorasi Data

Analisis dimulai dari eksplorasi data yang bertujuan untuk melihat plot di antara variabel bebas dan variabel terikat, serta korelasi antar variabel yang diduga secara signifikan mempengaruhinya. Melakukan eksplorasi data berdasarkan plot antara variabel terikat (Y) dengan setiap variabel bebasnya (X). Analisis hasil plot disimpulkan bahwa titik menyebar cenderung mengikuti garis lurus yang artinya hubungan linier terjadi di antara



variabel terikat (Y) dan variabel bebasnya (X).

GAMBAR 1. Scatterplot antara Variabel Terikat dan Variabel Bebas

Koefisien korelasi bertujuan untuk mengetahui besarnya hubungan dari setiap variabel yang digunakan. Hasil yang diperoleh dari penghitungan nilai koefisien korelasi untuk setiap variabel yang digunakan terdapat pada lampiran 1. Terlihat pada lampiran 1 bahwa lebih dari satu variabel bebas yang sangat berkorelasi tinggi dengan variabel bebas yang lain, yaitu variabel X4, X5, X6 dan X7 yang memiliki nilai korelasi kuat mendekati angka 1 atau korelasi sempurna. Sehingga disimpulkan bahwa terdapat masalah multikolinieritas pada data stunting yang digunakan.

Statistika deskriptif

Berikut adalah gambaran umum data penelitian yang diringkas dalam statistika deskriptif sebagai berikut:

TABEL 1. Statistik Deskriptif Variabel Terikat dan Variabel Bebas

Variabel	Rata - rata	Standar Deviasi	Variansi	Minimum	Maksimum
Y	30,25	5,302	28,11	17,60	42,60
X1	45,43	10,63	112,9	25,69	64,28
X2	60,53	5,99	35,89	44,04	72,49
X3	79,54	13,46	181,1	42,00	99,50
X4	84,81	10,07	101,4	50,70	99,20
X5	59,80	15,50	240,3	22,00	91,00
X6	66,24	15,13	229,1	25,90	96,00
X7	76,77	11,15	124,3	37,90	96,90
X8	30,94	7,86	61,80	15,50	55,30
X9	6,300	1,476	2,180	3,200	8,900
X10	6,529	3,480	12,108	2,100	14,000

X11 72,95 9,25 85,56 49,37 90,90

Regresi Linier Berganda dengan Metode Kuadrat Terkecil

Berdasarkan faktor-faktor yang mempengaruhi *stunting* di Indonesia diperoleh model regresi linear berganda menggunakan MKT dengan menggunakan software R sebagai berikut:

$$Y = 30,3 + 0,1411 X1 - 0,061 X2 + 0,066 X3 - 0,431 X4 - 0,381 X5 + 0,318 X6 + 0,201 X7 + 0,3023 X8 + 0,463 X9 + 0,107 X10 + 0,0257 X11$$

Nilai F-hitung yang didapatkan dari analisis regresi linier berganda menggunakan metode kuadrat terkecil menghasilkan model sebesar 5,20 dan p-value sebesar 0,0005 sehingga dihasilkan model yang nyata pada taraf 5%. Hasil analisis ragam menggunakan metode kuadrat terkecil terlihat pada Tabel 2.

TABEL 2. Analisis Ragam Hasil Metode Kuadrat Terkecil

Keragaman	db	Jumlah Kuadrat	Kuadrat Tengah	F-hitung
Regresi	11	669,969	60,906	5,20
Sisaan	22	257,716	11,714	
Total	33	927,685		

Dari sebelas variabel bebas yang digunakan dalam analisis hanya persentase imunisasi DPT-HB (X5) dan tinggi badan ibu (X8) yang memiliki pengaruh nyata terhadap balita *stunting* di Indonesia (Tabel 3). Hal tersebut terlihat dari p-value yang kurang dari 5%.

TABEL 3. Koefisien Regresi Hasil Metode Kuadrat Terkecil

Variabel	Koefisien	t-value	p-value
Constant	30,3	2,52	0,020
X1	0,1411	1,98	0,060
X2	-0,061	-0,35	0,727
X3	0,066	0,43	0,673
X4	-0,431	-1,50	0,148
X5*	-0,381	-2,13	0,045
X6	0,318	1,25	0,225
X7	0,201	0,89	0,382
X8*	0,3023	3,19	0,004
X9	0,463	0,98	0,339
X10	0,107	0,55	0,590
X11	0,0257	0,31	0,761

Ket : *) Variabel yang signifikan

Pendeteksian Multikolinieritas

Multikolinieritas dapat diidentifikasi menggunakan nilai Variance Inflation Factor (VIF), berikut adalah nilai VIF dari masing-masing variabel bebas:

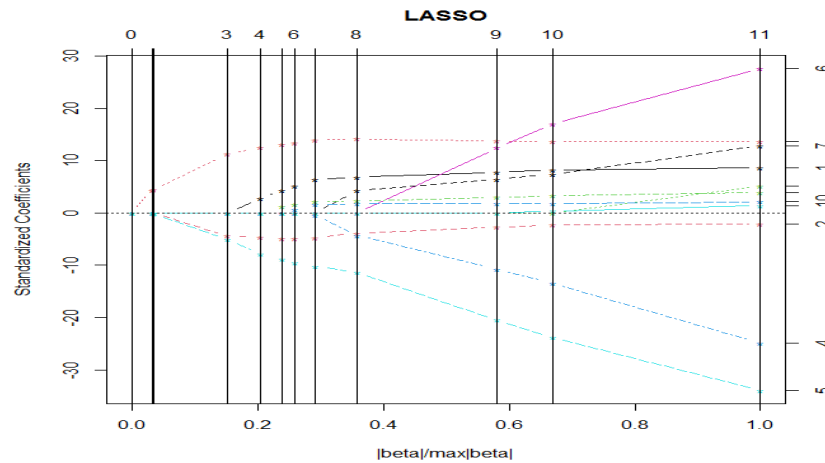
TABEL 4. Nilai VIF untuk Setiap Variabel Bebas

Variabel Bebas	VIF
X1	1,61
X2	2,99
X3	12,20
X4	23,51
X5	21,72
X6	41,84
X7	17,69
X8	1,56
X9	1,38
X10	1,31
X11	1,68

Pada Tabel 4 diketahui bahwa lima variabel bebas yang memiliki nilai VIF lebih besar dari 10 adalah variabel bebas Imunisasi Hepatitis B (X3), Imunisasi *Bacillus Calmette-Guerin* (X4), Imunisasi DPT-HB (X5), Imunisasi Polio (X6) dan Imunisasi Campak (X7) sehingga dapat disimpulkan bahwa terdapat masalah multikolinieritas. Korelasi yang kuat diantara X3 dan X4 dengan nilai korelasi sebesar 0,908 dan korelasi lainnya diantara variabel X5,X6,X7 yang mengindikasikan terjadinya masalah multikolinieritas antara variabel tersebut. Selain itu, model regresi penduga dengan MKT menghasilkan nilai koefisien determinasi yang cukup besar, dengan nilai R^2 yang berarti besarnya pengaruh ke sebelas variabel bebas terhadap variabel terikat sebesar 72,22% sedangkan sisanya yakni 27,78% dipengaruhi faktor-faktor lain di luar model. Nilai R-sq (adj) yang diperoleh sebesar 58,33% berarti bahwa variasi keragaman *stunting* yang mampu dijelaskan oleh variabel bebas dalam model sebesar 58,33% dan 41,67 % sisanya dijelaskan oleh variabel lain diluar model.

Pendugaan Koefisien Menggunakan LASSO

Masalah multikolinieritas sering terjadi pada analisis regresi linier berganda dan dapat diatasi salah satunya dengan menggunakan metode LASSO. Dalam penelitian ini, untuk dapat menyelesaikan regresi LASSO digunakan software RStudio dan dibantu dengan menggunakan package LAR. Dalam penyelesaian regresi LASSO ini digunakan pendekatan numerik menggunakan algoritma LAR. Berikut plot tahapan variabel bebas yang masuk ke dalam model:



GAMBAR 2. Plot yang di Hasilkan oleh Algoritma LAR Sebagai Penduga Koefisien LASSO

Gambar 2 adalah hasil modifikasi algoritma LAR untuk LASSO. Secara bertahap dilakukan penduga koefisien LASSO dengan ditetapkan semua koefisien awal bernilai nol. Kemudian, variabel bebas yang memiliki korelasi tinggi dengan galat secara bertahap masuk ke dalam model.

TABEL 5. Variabel Bebas yang Masuk dalam Model untuk Setiap Tahapan pada Metode LASSO

Tahapan	Variabel bebas yang masuk ke model
1	X8
2	X8 X5
3	X8 X5 X2
4	X8 X5 X2 X1
5	X8 X5 X2 X1 X9
6	X8 X5 X2 X1 X9 X10
7	X8 X5 X2 X1 X9 X10 X4
8	X8 X5 X2 X1 X9 X10 X4 X7
9	X8 X5 X2 X1 X9 X10 X4 X7 X6
10	X8 X5 X2 X1 X9 X10 X4 X7 X6 X11
11	X8 X5 X2 X1 X9 X10 X4 X7 X6 X11 X3

Pada iterasi pertama, variabel X8 merupakan variabel yang berkorelasi paling tinggi dengan galat dibandingkan dengan variabel lainnya sampai $\sum |\hat{\beta}_j| / \max \sum |\hat{\beta}_j|$ sekitar 0,04045 masuk variabel X5. Kemudian variabel X2 masuk ke dalam model dengan $\sum |\hat{\beta}_j| / \max \sum |\hat{\beta}_j|$ sekitar 0,17381 sampai variabel X3 yang merupakan variabel terakhir

masuk ke dalam model (Tabel 5). Setiap tahapan koefisien model regresi dapat dilihat pada Lampiran 2. Nilai $\sum |\hat{\beta}_j| / \max \sum |\hat{\beta}_j|$ untuk setiap tahapan terdapat di lampiran 3.

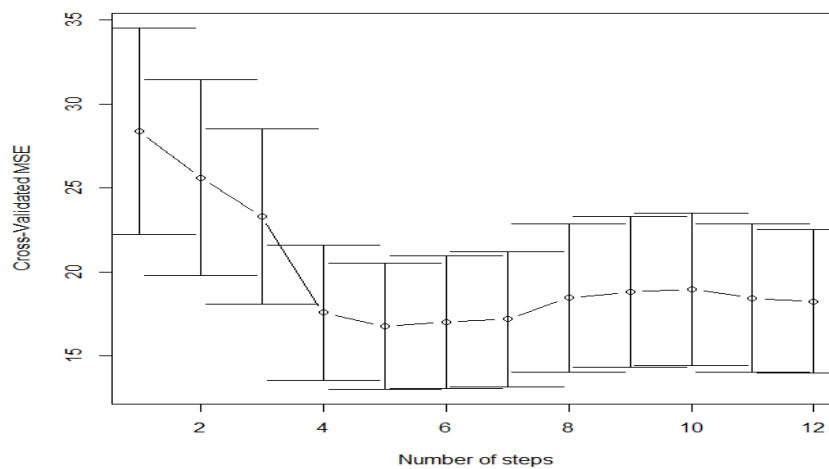
Pemilihan Model Terbaik Metode LASSO

Dalam pemilihan model terbaik LASSO dilakukan proses validasi silang k-fold dengan menggunakan mode step. Mode step dilakukan dengan menghitung nilai dari validasi silang untuk setiap tahapan dimana satu variabel masuk ke dalam model. Nilai s adalah nilai validasi silang yang dihitung berdasarkan $\sum |\hat{\beta}_j| / \max \sum |\hat{\beta}_j|$. Pemilihan model terbaik dalam metode LASSO ditentukan berdasarkan nilai dari s yang meminimumkan CV MSE (Tabel 6).

TABEL 6. Nilai CV MSE yang diperoleh dari Mode Step

Tahap	CV MSE
1	6,166708
2	5,833397
3	5,232610
4	4,030932
5	3,767157
6	3,935303
7	4,024921
8	4,444235
9	4,505852
10	4,548136
11	4,427273
12	4,269151

Pemilihan model terbaik menggunakan mode step menghasilkan nilai CV MSE model terbaik ada pada tahap 5 karena memperoleh nilai CV MSE minimum sebesar 3,767157. Dapat dilihat juga pada gambar 3 bahwa pada tahap 5 memberikan nilai CV yang merupakan titik belok dari turun menjadi naik pertama.



GAMBAR 3. Mode Step Menghasilkan Nilai Validasi Silang

Metode validasi silang dengan mode step mendapatkan model terbaik LASSO yang terpilih adalah model pada tahap ke-5. Sehingga lima variabel bebas yang masuk ke dalam model pada tahap ke-5 adalah asi eksklusif (X1), konsumsi protein (X2), imunisasi DPT-HB (X5), tinggi badan ibu (X8) dan diare (X9).

Model Hasil LASSO

Koefisien regresi hasil analisis dengan menggunakan LASSO dapat terlihat pada Tabel berikut:

TABEL 7. Koefisien Regresi Hasil LASSO

Variabel	LASSO
Intersep	30,3
X1	0,071
X2	-0,145
X3	0,000
X4	0,000
X5	-0,099
X6	0,000
X7	0,000
X8	0,289
X9	0,141
X10	0,000
X11	0,000

Pada Tabel 7 terlihat bahwa regresi LASSO menyusutkan koefisien hasil MKT sampai tepat nol kemudian secara otomatis variabel-variabel bebas akan diseleksi dari model yang terpilih. Nilai dari validasi silang minimum pada pemilihan model terbaik dengan mode step adalah $s = 0,29836$.

Berikut model persamaan regresi LASSO menggunakan algoritma LAR:

$$\hat{Y} = 30,3 + 0,071 X1 - 0,145 X2 - 0,099 X5 + 0,289 X8 + 0,141 X9$$

Berdasarkan model LASSO yang diperoleh, disimpulkan bahwa variabel bebas yang berpengaruh nyata terhadap variabel terikat yaitu *stunting* adalah variabel asi eksklusif (X1), konsumsi protein (X2), imunisasi DPT-HB (X5), tinggi badan ibu (X8) dan diare (X9). Sedangkan variabel-variabel bebas yang terseleksi dari model LASSO adalah variabel bebas yang tidak signifikan secara nyata terhadap *stunting*. Sehingga, metode LASSO dapat menyederhanakan model dan secara tidak langsung dapat mengatasi masalah multikolinearitas.

KESIMPULAN DAN SARAN

Pemilihan model terbaik untuk metode LASSO menggunakan proses validasi silang adalah pada $s = 0,29836$ dengan model sebagai berikut:

$$\hat{Y} = 0,071 X1 - 0,145 X2 - 0,099 X5 + 0,289 X8 + 0,141 X9$$

Berdasarkan model LASSO yang diperoleh, disimpulkan bahwa variabel bebas yang berpengaruh nyata terhadap variabel stunting adalah variabel asi eksklusif (X1), konsumsi protein (X2), imunisasi DPT-HB (X5), tinggi badan ibu (X8) dan diare (X9). Saran yang dapat dilakukan untuk penelitian kedepannya adalah membandingkan metode kuadrat terkecil, ridge, dan LASSO untuk dapat mengetahui sifat-sifat dari ketiga metode.

REFERENSI

- Aridiah Okky, Farah. 2015. "Faktor-Faktor Yang Mempengaruhi Kejadian Stunting Pada Anak Balita di Wilayah Pedesaan Dan Perkotaan (The Factors Affecting Stunting on Toddlers in Rural and Urban Areas)." e-Jurnal Pustaka Kesehatan 3 No 1.
- Dewanti, C., Ratnasari, V., & Rumiati, T. (2019). Pemodelan Faktor-faktor yang Memengaruhi Status Balita Stunting di Provinsi Jawa Timur Menggunakan Regresi Probit Biner. *Jurnal Sains Dan Seni Its*, 8(2). http://ejurnal.its.ac.id/index.php/sains_seni/article/view/48519
- Dewi, Y. S. (2010). OLS, LASSO dan PLS Pada data Mengandung Multikolinieritas. *Jurnal Ilmu Dasar*, 11(1), 83–91.
- Draper, N. R. dan Smith, H. 1992. Analisis Regresi Terapan. Ed. ke-2. Terjemahan oleh Bambang Sumantri. Gramedia Pustaka Utama, Jakarta.
- Farmani, K., Kencana, P., & Sukarsa, K. (2012). Perbandingan Analisis Least Absolute Shrinkage And Selection Operator Dan Partial Least Squares (Studi Kasus: Data Microarray). *E-Jurnal Matematika*, 1(1), 75–80. <https://doi.org/10.24843/MTK.2012.v01.i01.p013>
- Girma W, Ganebo T. 2002. Determinants of Nutritional Status of Woman and Children in Ethiopia. Calverton, Maryland, USA: ORC Macro
- Hastie, T., Tibshirani, R, Friedman, J. 2008. The Elements of Statistical Learning Data Mining, Inference, and Prediction Second Edition. New York: Springer.
- Hoerl AE, Kennard RW. 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* 12(1): 55-67.
- Hoffman DJ, Sawaya AL, Verreschi I, Tucker KL, Roberts SB, 2000. Why are nutritionally stunted children at increased risk of obesity? Studies of metabolic rate and fat oxidation in shantytown children from São Paulo, Brazil. *Am J Clin Nutrition* 72:702–7
- Montgomery, D., & Peck, E. (1992). Introduction Linear Regression Analysis 2nd Edition. John Wiley and Sons.
- Nurfitri Imro'ah, I. S. N. N. D. (2020). Analisis Regresi Dengan Metode Least Absolute Shrinkage and Selection Operator (Lasso) Dalam Mengatasi Multikolinearitas. *Bimaster : Buletin Ilmiah Matematika, Statistika Dan Terapannya*, 9(1), 31–38. <https://doi.org/10.26418/bbimst.v9i1.38029>
- Picauly, I. (2013). Analisa Determinan dan Pengaruh Stunting Terhadap Prestasi Belajar Anak Sekolah di Kupang dan Sumba Timur NTT. Diperoleh Tanggal 10 Januari 2019. *Jurnal Gizi dan Pangan*.
- Prabowo, F., Wilandari, Y., & Rusgiyono, A. (2015). Pemodelan pertumbuhan ekonomi jawa tengah menggunakan pendekatan least absolute shrinkage and selection operator (Lasso). *Jurnal Gaussian*, 4(4), 855–864.
- Riskesdas, K. (2018). Hasil Utama Riset Kesehata Dasar (RISKESDAS). *Journal of Physics A: Mathematical and Theoretical*, 44(8), 1–200. <https://doi.org/10.1088/1751-8113/44/8/085201>

- Setiawan, E., dkk. (2018). Faktor-Faktor yang Berhubungan Dengan Kejadian Stunting Pada Anak Usia 24-59 Bulan di Wilayah Kerja Puskesmas Andalas Kecamatan Padang Timur Kota Padang Tahun 2018. *Jurnal Kesehatan Andalas*, 7(2), 275-284.
- Soleh, A. M. dan Aunuddin. 2013. LASSO: Solusi Alternatif Seleksi Peubah dan Penyusutan Koefisien Model Regresi. *Indonesian Journal of Statistics*. 18 (1): 21-27.
- Tibshirani, R. (1996). Regression Shrinkage and Selection via The LASSO. *Journal Royal Statistic Soc B.*, Volume 58 No. 1, 267 - 288.
- Tibshirani, R. (2017). *Sparsity, The Lasso, and Friends*. Springer
- Unicef, 1990. *Strategy for improved nutrition of children and women in developing countries*. New York
- Santi, V. M, Notodiputro, K. A, Sartono, B. (2019). *Variable Selection Methods for Generalized Linear Mixed Models by Penalized Lasso in Modeling the Scores of Indonesian Students in Programme for International Student Assessment (PISA)*.
- WHO. 2010. *Nutrition Landscape Information System (NLIS): Country Profile Indicators InterpretationGuide*. [online].