

SKRIPSI

**OPTIMISASI REGRESI LOGISTIK LASSO DENGAN
KRITERIA AKAIKE INFORMATION CRITERION (AIC)
UNTUK KLASIFIKASI FAKTOR RISIKO STUNTING PADA
ANAK BERDASARKAN DATA SSGI 2024**



SHUN NAFINGATUN NURUL UMMAH
22106010028
SUNAN KALIJAGA
YOGYAKARTA

PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA

2026

**OPTIMISASI REGRESI LOGISTIK LASSO DENGAN
KRITERIA AKAIKE INFORMATION CRITERION (AIC)
UNTUK KLASIFIKASI FAKTOR RISIKO STUNTING PADA
ANAK BERDASARKAN DATA SSGI 2024**

Skripsi

Untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana S-1
Program Studi Matematika



diajukan oleh

SHUN NAFINGATUN NURUL UMMAH

22106010028

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Kepada

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA**

2026



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-1874/Un.02/DST/PP.00.9/08/2026

Tugas Akhir dengan judul : Optimisasi Regresi Logistik LASSO dengan Kriteria Akaike Information Criterion (AIC) untuk Klasifikasi Faktor Risiko Stunting pada Anak Berdasarkan Data SSGI 2024

yang dipersiapkan dan disusun oleh:

Nama : SHUN NAFINGATUN NURUL UMMAH
Nomor Induk Mahasiswa : 22106010028
Telah diujikan pada : Jumat, 07 Agustus 2026
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Sri Utami Zuliana, S.Si., M.Sc., Ph.D.
SIGNED

Valid ID: 6a7e982bec0ec



Penguji I

Lilih Deva Martias, M.Sc.
SIGNED

Valid ID: 6a7d81a9e8616



Penguji II

Noor Saif Muhammad Mussafi, S.Si., M.Sc.,
Ph.D.
SIGNED

Valid ID: 6a7d3eb04d806



Yogyakarta, 07 Agustus 2026
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 6a83d83bee8a1



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

Assalamu 'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Shun Nafingatun Nurul Ummah

NIM : 22106010028

Judul Skripsi : Optimisasi Regresi Logistik LASSO dengan Kriteria Akaike Information Criterion (AIC) untuk Klasifikasi Faktor Risiko Stunting pada Anak Berdasarkan Data SSGI 2024

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Matematika.

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu 'alaikum wr. wb.

Yogyakarta, 23 Juli 2026

Pembimbing

Sri Utami Zuliana, S.Si., M.Sc., Ph.D.

NIP. 19741003 200003 2 002

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Shun Nafingatun Nurul Ummah
NIM : 22106010028
Program Studi : Matematika
Fakultas : Sains dan Teknologi

Dengan ini menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sesungguhnya skripsi ini merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi atau saduran dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 23 Juli 2026



Shun Nafingatun Nurul Ummah

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

SKRIPSI INI SAYA PERSEMBAHKAN SEBAGAI BENTUK TANDA CINTA DAN
TERIMA KASIH YANG MENDALAM KEPADA ORANG TUA, KAKAK, DAN
KELUARGA BESAR YANG SELALU MEMBERIKAN DUKUNGAN, SEMANGAT,
SERTA DOA YANG TAK HENTI MENGIRINGI PROSES PENYELESAIAN
SKRIPSI INI.

TAK LUPA, PERSEMBAHAN INI JUGA UNTUK **ALMAMATER TERCINTA**
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA YOGYAKARTA
YANG TELAH MENJADI RUANG TUMBUH DAN BELAJAR SELAMA MASA
STUDI.

"MATEMATIKA ANGKATAN 2022 TERCINTA"

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

MOTTO

"Tidak akan menimpa kami melainkan apa yang telah ditetapkan Allah bagi kami."

(QS. At Taubah:51)

"Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya."

(QS. Al Baqarah:286)

"Hidup bukanlah tentang memaksakan jawaban, tetapi belajar untuk hidup berdampingan dengan ketidakpastian."

"Setiap perjalanan pasti akan berakhir. Jangan terburu-buru."

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat dan karunia-Nya yang tak ternilai, berupa keimanan, kesabaran, kekuatan, serta kemudahan sehingga penulis dapat menyelesaikan skripsi ini. Shalawat dan salam semoga senantiasa tercurah kepada Nabi Muhammad SAW. Atas izin-Nya pula, skripsi yang berjudul “Optimisasi Regresi Logistik LASSO dengan Kriteria *Akaike Information Criterion* (AIC) untuk Klasifikasi Faktor Risiko Stunting pada Anak Berdasarkan Data SSGI 2024” dapat diselesaikan dengan baik.

Penulis menyadari bahwa penyusunan skripsi ini tidak terlepas dari berbagai kendala dan tantangan. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada seluruh pihak yang telah memberikan bantuan, dukungan, serta motivasi selama proses penyusunan skripsi ini hingga dapat diselesaikan.

Dengan penuh ketulusan, penulis menyampaikan rasa terima kasih dan penghargaan kepada:

1. Bapak Prof. Noorhaidi Hasan, S.Ag., M.A., M.Phil., Ph.D., selaku Rektor Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
2. Ibu Prof. Dr. Dra. Hj. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
3. Ibu Dr. Epha Diana Supandi, S.Si., M.Sc., selaku Ketua Program Studi Matematika.
4. Ibu Sri Utami Zuliana, S.Si., M.Sc., Ph.D., selaku dosen pembimbing, atas kesediaan waktu, bimbingan, arahan, serta ilmu yang telah diberikan selama

proses penyusunan skripsi ini. Petunjuk dan masukan yang diberikan sangat membantu penulis dalam menyelesaikan penelitian ini dan menjadi bekal berharga untuk perjalanan akademik ke depan. Penulis juga menyampaikan permohonan maaf atas segala kekhilafan yang mungkin terjadi selama proses bimbingan.

5. Seluruh dosen Program Studi Matematika serta staf Fakultas Sains dan Teknologi yang telah memberikan ilmu pengetahuan, pelayanan, dan dukungan kepada penulis sejak awal hingga akhir masa perkuliahan.
6. Orang tua, kakak, serta keluarga besar penulis yang senantiasa memberikan nasihat, kasih sayang, doa, dan dukungan moral maupun material sehingga penulis dapat menyelesaikan studi ini.
7. Keluarga besar Mahasiswa Matematika angkatan 2022 yang telah kebersamaan penulis selama empat tahun masa perkuliahan, saling mendukung, berbagi ilmu, serta menciptakan kenangan berharga.
8. Teman-teman seperbimbingan, terima kasih atas dukungan, semangat, diskusi, serta kebersamaan yang terjalin selama proses penyusunan skripsi ini.
9. Sahabat-sahabat penulis yang selalu hadir dalam suka dan duka, memberikan dukungan moral, doa, serta semangat yang tak ternilai selama proses perkuliahan hingga penyusunan skripsi ini.
10. Seluruh pihak yang telah memberikan bantuan, dukungan, serta doa kepada penulis, baik secara langsung maupun tidak langsung, yang tidak dapat disebutkan satu per satu. Semoga segala kebaikan yang telah diberikan mendapat balasan berupa keberkahan dan kebahagiaan.

Penulis menyadari sepenuhnya bahwa skripsi ini masih memiliki keterbatasan dan belum sempurna. Oleh karena itu, penulis dengan terbuka menerima segala saran dan kritik yang bersifat membangun demi penyempurnaan karya ini. Penulis berharap skripsi ini dapat memberikan manfaat, khususnya bagi penulis, serta bagi berbagai pihak yang membutuhkan sebagai bahan referensi dalam penulisan dan pengembangan keilmuan. Akhir kata, semoga Allah SWT senantiasa melimpahkan rahmat dan hidayah-Nya kepada kita semua. Aamiin.

Yogyakarta, 20 Juli 2026

Shun Nafingatun



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PERNYATAAN KEASLIAN	iv
HALAMAN PERSEMBAHAN	v
MOTTO	vi
PRAKATA	vii
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMBANG	xv
ABSTRAK	.xviii
ABSTRACT	xix
I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah	4
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
1.6. Tinjauan Pustaka	6
1.7. Sistematika Penulisan	11
II LANDASAN TEORI	13
2.1. Model Linear	13

2.2. Multikolinearitas	14
2.3. Regresi Logistik	15
2.3.1. Model Regresi Logistik	16
2.3.2. Estimasi Parameter dengan Maximum Likelihood Estima- tion (MLE)	17
2.3.3. Algoritma <i>Iteratively Reweighted Least Squares</i> (IRLS)	18
2.4. LASSO	27
2.4.1. Formulasi LASSO	27
2.4.2. Regresi Logistik LASSO	28
2.4.3. Algoritma IRLS dan <i>Coordinate Descent</i> untuk Regresi Lo- gistik LASSO	29
2.5. Akaike Information Criterion (AIC)	31
2.5.1. Definisi dan Formulasi AIC	31
2.5.2. Interpretasi AIC	32
2.5.3. AIC dalam LASSO untuk Regresi Logistik	32
2.6. Stunting dan Klasifikasi Status Stunting	34
2.6.1. Prevalensi Stunting dan Target Nasional	34
2.6.2. Klasifikasi Status Stunting Tinggi dan Rendah	34
2.6.3. Faktor-Faktor Risiko Stunting	35
III METODE PENELITIAN	36
3.1. Jenis Penelitian	36
3.2. Sumber Data	36
3.2.1. Populasi dan Sampel	37
3.3. Variabel Penelitian	38
3.3.1. Variabel Respons	38
3.3.2. Variabel Prediktor	38

3.4. Metode Pengumpulan Data	40
3.5. Software	41
3.6. Source code	41
3.7. Tahapan Analisis	42
3.8. Diagram Alur Penelitian	45
IV HASIL DAN PEMBAHASAN	46
4.1. Deskripsi Data Penelitian	46
4.2. Analisis Deskriptif	48
4.3. Uji Multikolinearitas	50
4.4. Persiapan Data dan Standardisasi	53
4.5. Pembentukan Model Regresi Logistik Lasso	56
4.6. Penentuan Parameter Optimum Menggunakan AIC	59
4.7. Hasil Regresi Logistik LASSO dengan AIC	62
4.8. Interpretasi Odds Ratio	63
4.9. Evaluasi Model	66
4.10. Perbandingan Regresi Logistik Penuh dan Regresi Logistik LASSO	
AIC	67
V PENUTUP	69
5.1. Kesimpulan	69
5.2. Saran	71
DAFTAR PUSTAKA	73
LAMPIRAN	76
A Data Penelitian	77
B Source Code Program R	80

DAFTAR TABEL

3.1 Variabel Penelitian	40
4.1 Pengelompokan Provinsi Berdasarkan Kategori Stunting	47
4.2 Statistik Deskriptif Variabel Penelitian	49
4.3 Hasil Uji <i>Variance Inflation Factor</i> (VIF)	51
4.4 Rata-rata dan Standar Deviasi Variabel Prediktor Sebelum Standarisasi	55
4.5 Hasil Verifikasi Standardisasi Variabel Prediktor	55
4.6 Koefisien Regresi Logistik LASSO	57
4.7 Hasil Penentuan Parameter Optimum	60
4.8 Koefisien Model Regresi Logistik LASSO	62
4.9 Odds Ratio Variabel Terpilih	63
4.10 Confusion Matrix Model LASSO	66
4.11 Perbandingan Model Regresi Logistik Penuh dan Regresi Logistik LASSO AIC	67

DAFTAR GAMBAR

4.1 Nilai VIF Variabel Prediktor	52
4.2 Lintasan Koefisien Regresi Logistik LASSO	58
4.3 Kurva AIC terhadap Nilai Log Lambda	61
4.4 Odds Ratio Variabel Model LASSO-AIC	64
4.5 Kurva ROC Model LASSO-AIC	66

DAFTAR LAMBANG

Y	= variabel respons (0 = stunting rendah, 1 = stunting tinggi)
X_j	= variabel prediktor ke- j
$\mathbf{X}$	= matriks desain berukuran $n \times (p + 1)$
n	= jumlah observasi (provinsi)
p	= jumlah variabel prediktor
β_0	= intersep (konstanta) dalam model regresi
β_j	= koefisien regresi untuk variabel prediktor ke- j
$\boldsymbol{\beta}$	= vektor koefisien regresi $(\beta_0, \beta_1, \dots, \beta_p)^T$
ε	= error term (galat) dalam model linear
σ^2	= varians dari error term
$E(Y \mathbf{X})$	= nilai harapan variabel respons bersyarat pada $\mathbf{X}$
$\ell(\boldsymbol{\beta})$	= fungsi log-likelihood
$L(\boldsymbol{\beta})$	= fungsi likelihood
$p(\mathbf{X})$	= probabilitas $Y = 1$ bersyarat pada $\mathbf{X}$, $P(Y = 1 \mathbf{X})$
p_i	= probabilitas $Y_i = 1$ untuk observasi ke- i
η_i	= linear predictor untuk observasi ke- i
$\mathbf{x}_i$	= vektor variabel prediktor untuk observasi ke- i $(1, x_{i1}, \dots, x_{ip})^T$
$g(\cdot)$	= fungsi penghubung (link function)
μ_i	= nilai harapan variabel respons untuk observasi ke- i , $E(Y_i \mathbf{x}_i)$
$\hat{\boldsymbol{\beta}}_{MLE}$	= estimator *Maximum Likelihood Estimation* (MLE)

t	= indeks iterasi pada algoritma IRLS
$\beta^{(t)}$	= vektor koefisien pada iterasi ke- t
$\eta_i^{(t)}$	= linear predictor pada iterasi ke- t untuk observasi ke- i
$p_i^{(t)}$	= probabilitas prediksi pada iterasi ke- t untuk observasi ke- i
$w_i^{(t)}$	= bobot untuk observasi ke- i pada iterasi ke- t
$\mathbf{W}^{(t)}$	= matriks bobot diagonal pada iterasi ke- t
$z_i^{(t)}$	= working response untuk observasi ke- i pada iterasi ke- t
$\mathbf{z}^{(t)}$	= vektor working response pada iterasi ke- t
$\mathbf{y}$	= vektor variabel respons $(y_1, y_2, \dots, y_n)^T$
$\mathbf{p}^{(t)}$	= vektor probabilitas prediksi pada iterasi ke- t
$H(\beta)$	= matriks Hessian dari fungsi log-likelihood
$\nabla \ell(\beta)$	= vektor gradien dari fungsi log-likelihood
$\ \cdot\ _2$	= norma Euclidean (L_2)
ε	= nilai toleransi konvergensi (IRLS)
λ	= parameter penalti pada LASSO ($\lambda \geq 0$)
$\hat{\beta}_{LASSO}$	= estimator LASSO
OR_j	= *odds ratio* untuk variabel prediktor ke- j
Odds	= rasio odds $\frac{p}{1-p}$
$S(a, \lambda)$	= fungsi soft-thresholding
$(a)_+$	= fungsi positif: $\max(a, 0)$
$\tilde{x}_{ij}$	= nilai variabel prediktor ke- j observasi ke- i setelah standarisasi
$\bar{x}_j$	= rata-rata sampel untuk variabel ke- j
s_j	= standar deviasi sampel untuk variabel ke- j
$r_{i(-j)}$	= residual parsial untuk observasi ke- i tanpa variabel ke- j
c_j	= korelasi berbobot untuk variabel ke- j

- ε_{cd} = toleransi konvergensi untuk coordinate descent
 ε_{irls} = toleransi konvergensi untuk IRLS
AIC = Akaike Information Criterion
 k = jumlah parameter dalam model
 $df(\lambda)$ = derajat bebas efektif (*effective degrees of freedom*) model LASSO
 $\hat{\theta}_{MLE}$ = estimator MLE untuk parameter θ
 ΔAIC = selisih nilai AIC antara suatu model dengan model terbaik
 Z_i = nilai variabel setelah distandardisasi
 X_i = nilai asli variabel pada observasi ke- i
 $\bar{X}$ = rata-rata variabel sebelum standardisasi
 S_X = standar deviasi variabel sebelum standardisasi.

ABSTRAK

Optimisasi Regresi Logistik LASSO dengan Kriteria Akaike Information Criterion (AIC) untuk Klasifikasi Faktor Risiko Stunting pada Anak Berdasarkan Data SSGI 2024

Oleh

Shun Nafingatun Nurul Ummah

22106010028

Dalam regresi logistik biner, multikolinearitas antarvariabel prediktor dapat menyebabkan estimasi parameter menjadi tidak stabil sehingga menurunkan kualitas model. Untuk mengatasi permasalahan tersebut, digunakan regresi logistik LASSO (*Least Absolute Shrinkage and Selection Operator*) yang menerapkan penalti norma- L_1 sehingga mampu melakukan seleksi variabel dan penyusutan koefisien secara simultan. Penentuan parameter penalti λ dilakukan menggunakan *Akaike Information Criterion* (AIC) untuk memperoleh model yang memberikan keseimbangan antara kecocokan model dan kompleksitasnya. Metode tersebut diterapkan pada data *Survei Status Gizi Indonesia* (SSGI) Tahun 2024 yang mencakup 38 provinsi di Indonesia untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap status stunting. Hasil penelitian menunjukkan bahwa regresi logistik LASSO dengan kriteria AIC menghasilkan model yang lebih sederhana dengan tujuh variabel prediktor, yaitu Jarak Kehamilan Risiko, Berat Badan Lahir Rendah (BBLR), ASI Eksklusif, *Minimum Dietary Diversity* (MDD), Imunisasi Lengkap, ISPA, dan Pendidikan Ibu Tinggi. Model memiliki nilai AIC lebih rendah dibandingkan regresi logistik penuh, dengan AUC sebesar 1,00 dan akurasi klasifikasi sebesar 100%. Hasil tersebut menunjukkan bahwa model LASSO dengan kriteria AIC mampu menghasilkan model yang parsimoni dengan kemampuan klasifikasi yang sangat baik.

Kata kunci: Regresi logistik LASSO, AIC, Stunting, SSGI.

ABSTRACT

LASSO Logistic Regression Optimization with AIC for Stunting

Classification in Children Based on SSGI 2024 Data

By

Shun Nafingatun Nurul Ummah

22106010028

In binary logistic regression, multicollinearity among predictor variables can cause parameter estimates to become unstable, thereby reducing the quality of the model. To address this issue, LASSO (Least Absolute Shrinkage and Selection Operator) logistic regression is used, which applies an L_1 norm penalty, thereby enabling simultaneous variable selection and coefficient shrinkage. The penalty parameter λ was determined using the *Akaike Information Criterion* (AIC) to obtain a model that strikes a balance between model fit and complexity. This method was applied to data from the 2024 *Indonesian Nutrition Status Survey* (SSGI), which covers 38 provinces in Indonesia, to identify factors influencing stunting status. The results of the study show that LASSO logistic regression using the AIC criterion produces a simpler model with seven predictor variables: Interpregnancy Interval, Low Birth Weight (LBW), Exclusive Breastfeeding, *Minimum Dietary Diversity* (MDD), Complete Immunization, Acute Respiratory Infections (ARI), and High Maternal Education. The model has a lower AIC value than full logistic regression, with an AUC of 1.00 and a classification accuracy of 100%. These results indicate that the LASSO model using the AIC criterion is capable of producing a parsimonious model with excellent classification performance.

Keywords: LASSO Logistic Regression, AIC, Stunting, SSGI.

BAB I

PENDAHULUAN

1.1. Latar Belakang

Regresi logistik merupakan salah satu metode statistik yang paling sering digunakan dalam analisis data kesehatan, khususnya untuk memodelkan variabel respons biner atau kategorikal. Metode ini memungkinkan peneliti untuk mengidentifikasi faktor-faktor risiko dan mengestimasi probabilitas terjadinya suatu kejadian berdasarkan variabel prediktor yang tersedia. Dalam konteks analisis data kesehatan yang kompleks, regresi logistik memberikan interpretasi yang mudah dipahami melalui *odds ratio* dan dapat menangani berbagai jenis variabel prediktor secara simultan (Hosmer et al., 2013).

Namun, dalam praktiknya, analisis regresi logistik sering menghadapi tantangan ketika berhadapan dengan data yang memiliki jumlah variabel prediktor yang besar (*high-dimensional data*). Permasalahan multikolinearitas, *overfitting*, dan kesulitan dalam interpretasi model menjadi kendala utama yang dapat menurunkan akurasi dan generalisasi model. *Overfitting* terjadi ketika model terlalu kompleks dan mengikuti noise dalam *data training*, sehingga performanya buruk pada data baru. Untuk mengatasi permasalahan ini, dikembangkan metode regularisasi yang menambahkan *penalty term* pada fungsi objective untuk mengontrol kompleksitas model (Hastie et al., 2009).

Least Absolute Shrinkage and Selection Operator (LASSO) merupakan salah satu metode regularisasi yang paling populer dan efektif untuk mengatasi per-

masalah *overfitting* dan seleksi variabel secara simultan. Lasso menambahkan penalty L1 (jumlah nilai absolut dari koefisien) pada fungsi likelihood regresi logistik, yang memiliki karakteristik unik yaitu dapat mengecilkan beberapa koefisien menjadi tepat nol. Dengan demikian, Lasso tidak hanya melakukan regularisasi tetapi juga secara otomatis melakukan seleksi variabel, menghasilkan model yang lebih parsimoni dan mudah diinterpretasikan (Tibshirani, 1996). Keunggulan Lasso dalam menangani multikolinearitas dan mengidentifikasi variabel yang benar-benar berpengaruh menjadikannya pilihan yang tepat untuk analisis data kesehatan yang kompleks.

Meskipun Lasso memiliki kemampuan seleksi variabel yang baik, penentuan parameter tuning (λ) yang optimal menjadi krusial untuk menghasilkan model terbaik. Parameter lambda mengontrol tingkat regularisasi, dimana nilai yang terlalu kecil akan menghasilkan model yang kompleks (mendekati regresi logistik biasa), sedangkan nilai yang terlalu besar akan menghasilkan model yang terlalu sederhana. *Akaike Information Criterion* (AIC) merupakan salah satu kriteria pemilihan model yang banyak digunakan karena mempertimbangkan *trade-off* antara *goodness of fit* dan kompleksitas model. AIC didasarkan pada teori informasi dan bertujuan meminimalkan kehilangan informasi ketika model digunakan untuk merepresentasikan proses yang menghasilkan data (Akaike, 1974). Kombinasi antara Lasso dan AIC dapat menghasilkan model yang optimal, dimana Lasso melakukan seleksi variabel dan AIC memastikan pemilihan model dengan keseimbangan terbaik antara bias dan varians.

Status gizi merupakan salah satu indikator penting dalam menilai kesehatan masyarakat, terutama dalam konteks kesehatan anak dan ibu. Masalah gizi, baik gizi kurang maupun gizi lebih, memiliki dampak jangka pendek dan jangka panjang

terhadap kualitas sumber daya manusia. Di Indonesia, permasalahan gizi masih menjadi tantangan utama dalam pembangunan kesehatan. Survei Status Gizi Indonesia (SSGI) merupakan survei nasional yang dilaksanakan secara berkala untuk memantau status gizi masyarakat Indonesia, khususnya balita, ibu hamil, dan kelompok rentan lainnya. Data SSGI tahun 2024 memberikan informasi komprehensif mengenai prevalensi *stunting*, *wasting*, *underweight*, dan *overweight* di berbagai wilayah Indonesia.

Data SSGI tahun 2024 mengandung banyak variabel yang berpotensi mempengaruhi status gizi, mulai dari karakteristik individu (usia, jenis kelamin), karakteristik rumah tangga (tingkat pendidikan orang tua, status ekonomi, jumlah anggota keluarga), hingga faktor lingkungan (akses air bersih, sanitasi, akses pelayanan kesehatan). Dengan kompleksitas data yang tinggi ini, diperlukan metode analisis yang tepat untuk mengidentifikasi faktor-faktor dominan yang mempengaruhi status gizi. Penggunaan regresi logistik Lasso yang dioptimasi dengan AIC diharapkan dapat menghasilkan model yang akurat, parsimoni, dan mudah diinterpretasikan untuk mendukung pengambilan kebijakan dalam program perbaikan gizi di Indonesia.

Penelitian ini bertujuan untuk mengaplikasikan regresi logistik Lasso dengan optimisasi AIC pada data SSGI 2024 guna mengidentifikasi faktor-faktor yang paling berpengaruh terhadap status gizi di Indonesia. Dengan menggunakan pendekatan ini, diharapkan dapat diperoleh model prediksi yang optimal dengan variabel prediktor yang relevan, sehingga dapat memberikan rekomendasi yang evidence-based untuk intervensi program gizi nasional.

1.2. Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka rumusan masalah dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana mekanisme seleksi variabel dan estimasi parameter pada regresi logistik LASSO?
2. Bagaimana langkah-langkah optimisasi pada model regresi logistik LASSO dengan menggunakan *Akaike Information Criterion* (AIC)?
3. Bagaimana penerapan optimisasi regresi logistik LASSO menggunakan *Akaike Information Criterion* (AIC) terhadap data faktor klasifikasi stunting di Indonesia Tahun 2024?

1.3. Batasan Masalah

Agar fokus pada sasaran yang diharapkan maka perlu dilakukan pembatasan masalah dalam penelitian ini sebagai berikut:

1. Metode klasifikasi yang digunakan adalah regresi logistik dengan regularisasi LASSO.
2. Pemilihan parameter regularisasi menggunakan *Akaike Information Criterion* (AIC).
3. Data yang digunakan adalah data sekunder Survei Status Gizi Indonesia (SSGI) tahun 2024 yang diterbitkan oleh Kementerian Kesehatan Republik Indonesia.
4. *Software* yang digunakan adalah R versi R Core Team (2025).

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Menjelaskan mekanisme estimasi parameter pada regresi logistik LASSO menggunakan algoritma *Iteratively Reweighted Least Squares* (IRLS).
2. Menentukan langkah-langkah optimasi parameter regularisasi pada regresi logistik LASSO menggunakan *Akaike Information Criterion* (AIC).
3. Menerapkan optimisasi Regresi Logistik LASSO menggunakan *Akaike Information Criterion* (AIC) terhadap data faktor Klasifikasi Stunting pada Anak di Indonesia Tahun 2024

1.5. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis maupun praktis. Secara teoritis, penelitian ini berkontribusi pada pengembangan metode seleksi variabel dalam analisis regresi logistik melalui kombinasi regularisasi Lasso dan kriteria *Akaike Information Criterion* (AIC). Pendekatan *regularized information criteria* yang digunakan dalam penelitian ini memperkaya literatur statistika, khususnya dalam penanganan data dengan jumlah variabel prediktor yang relatif besar dibandingkan jumlah observasi. Selain itu, penelitian ini memberikan ilustrasi tentang mekanisme estimasi parameter pada regresi logistik Lasso menggunakan algoritma *Iteratively Reweighted Least Squares* (IRLS) serta prosedur optimasi pemilihan parameter penalti berbasis AIC yang dapat menjadi referensi bagi penelitian-penelitian selanjutnya.

Secara praktis, penelitian ini menyediakan kerangka analisis yang dapat diadopsi oleh peneliti lain dalam mengidentifikasi faktor-faktor dominan pada berbagai

kasus klasifikasi, tidak terbatas pada bidang kesehatan tetapi juga dapat diterapkan pada bidang sosial, ekonomi, dan lainnya. Hasil penelitian ini juga memberikan contoh penerapan regresi logistik Lasso dengan optimasi AIC pada data riil Survei Status Gizi Indonesia (SSGI) 2024, sehingga dapat menjadi acuan dalam analisis data serupa. Lebih lanjut, identifikasi faktor-faktor risiko stunting di tingkat provinsi yang diperoleh dari penelitian ini diharapkan dapat menjadi bahan pertimbangan berbasis data dalam pengambilan kebijakan penanggulangan stunting di Indonesia.

1.6. Tinjauan Pustaka

[Safitri et al. \(2022\)](#) dalam artikelnya yang berjudul *Faktor-Faktor yang Mempengaruhi Kejadian Stunting terhadap Balita menggunakan Analisis Regresi Logistik* menganalisis faktor-faktor yang mempengaruhi kejadian stunting pada balita di Posyandu Kasih Ibu Puskesmas Airtiris, Kampar, menggunakan metode regresi logistik biner. Hasil penelitian menunjukkan bahwa status gizi berat badan berdasarkan umur merupakan faktor yang berpengaruh signifikan terhadap kejadian stunting, dengan nilai odds ratio sebesar 7,761 yang berarti balita dengan status gizi buruk memiliki risiko stunting 7,761 kali lebih besar dibandingkan balita dengan status gizi normal. Penelitian ini juga melaporkan bahwa ketepatan klasifikasi model regresi logistik biner mencapai 83,6%, yang mengindikasikan bahwa model yang terbentuk fit atau sudah sesuai.

[Arofah & Sofro \(2022\)](#) dalam artikelnya yang berjudul *Penerapan Regresi Logistik Multinomial pada Tingkat Depresi Lansia* menganalisis faktor-faktor yang mempengaruhi tingkat depresi pada lansia menggunakan metode regresi logistik multinomial. Penelitian ini menggunakan *Akaike Information Criterion* (AIC) untuk menentukan model terbaik, di mana model dengan nilai AIC terkecil (109,836) dipilih sebagai model terbaik yang terdiri atas variabel usia, jenis kelamin, status

pernikahan, status pekerjaan, tingkat pendidikan, dan riwayat penyakit. Hasil penelitian menunjukkan bahwa usia merupakan faktor yang paling mempengaruhi tingkat depresi, dengan nilai odds ratio pada depresi ringan sebesar 1,372 dan depresi sedang sebesar 2,286.

[Sari et al.](#) (2026) dalam artikelnya yang berjudul *Pemodelan Spasial Kerawanan Longsor di Kabupaten Temanggung Menggunakan Regresi Logistik dengan Seleksi Variabel Stepwise AIC* menerapkan prosedur stepwise berbasis *Akaike Information Criterion* (AIC) pada regresi logistik untuk pemodelan kerawanan longsor. Hasil penelitian menunjukkan bahwa prosedur stepwise AIC mampu menurunkan nilai AIC secara bertahap dari 1267,33 menjadi 1258,46 pada model terbaik yang terdiri atas variabel elevasi, aspek lereng (cos), penggunaan lahan, jarak ke jalan, *Stream Power Index* (SPI), dan *Topographic Wetness Index* (TWI). Penelitian ini juga melaporkan bahwa model memiliki performa prediktif yang baik dengan nilai AUC sebesar 0,703.

[Rusdy et al.](#) (2025) dalam artikelnya yang berjudul *Determining Factors that Influence Unmet Need For Family Planning Using Geographically Weighted Logistic Regression With LASSO* mengembangkan metode *geographically weighted logistic regression* (GWLR) dengan Lasso untuk mengatasi masalah multikolinearitas pada data spasial. Penelitian ini menjelaskan bahwa GWLR mempertimbangkan faktor lokasi melalui fungsi pembobot, namun tidak dapat mengatasi multikolinearitas sehingga diperlukan LASSO yang menyusutkan nilai koefisien regresi hingga mendekati nol atau tepat nol. Hasil penelitian menunjukkan bahwa nilai AIC model GWLR dengan Lasso (31,918) lebih kecil dibandingkan model GWLR tanpa Lasso (38,879), yang mengindikasikan bahwa GWLR dengan Lasso lebih baik dalam memodelkan unmet need KB di Sulawesi Selatan.

Mkungudza et al. (2024) dalam artikelnya yang berjudul *Development of a diagnostic predictive model for determining child stunting in Malawi: a comparative analysis of variable selection approaches* membandingkan kinerja tujuh metode seleksi variabel, termasuk Lasso, dalam membangun model prediksi stunting di Malawi. Hasil penelitian menunjukkan bahwa LASSO efektif dalam menyeleksi variabel prediktor penting seperti indeks kekayaan rumah tangga, usia anak, dan berat badan lahir, dengan menghasilkan model yang parsimoni (sederhana) dan memiliki akurasi prediksi yang kompetitif dibandingkan metode lainnya. Penelitian ini menjadi bukti empiris bahwa Lasso sangat cocok untuk pemodelan faktor risiko stunting karena kemampuannya menangani banyak variabel prediktor yang saling berkorelasi.

Mataheru & Utari (2026) dalam artikelnya yang berjudul *Faktor Risiko yang Berhubungan dengan Kejadian Stunting pada Balita di Sulawesi Barat: Analisis Data SSGI 2024* menganalisis faktor risiko stunting pada balita di Provinsi Sulawesi Barat menggunakan desain potong lintang dengan analisis survei kompleks serta regresi logistik bivariat dan multivariat. Hasil penelitian menunjukkan bahwa prevalensi stunting di Sulawesi Barat sebesar 35,6%, dengan berat badan lahir rendah (BBLR) sebagai faktor risiko paling kuat (aOR = 2,08; 95% CI: 1,60-2,70). Faktor lain yang berhubungan signifikan meliputi jenis kelamin laki-laki (aOR = 1,24), akses sumber air minum tidak terlindungi (aOR = 1,43), pendidikan ibu rendah (aOR = 1,91), pendidikan ibu menengah (aOR = 1,36), serta riwayat kekurangan energi kronik (KEK) pada ibu saat hamil (aOR = 1,47).

Mochtar et al. (2024) dalam artikelnya yang berjudul *Analisis Regresi Logistik Biner pada Karakteristik Demografi Sosial Ekonomi dan Intervensi Gizi Spesifik Balita Stunting di Wilayah Kerja Puskesmas Jambula Kota Ternate* menganalisis

pengaruh karakteristik sosial ekonomi dan intervensi gizi spesifik terhadap kejadian stunting menggunakan regresi logistik biner. Hasil penelitian menunjukkan bahwa jumlah anggota keluarga ($OR = 0,842$, $p = 0,043$), jarak lahir anak ($OR = 9,3$, $p = 0,013$), sosial ekonomi keluarga ($OR = 11,8$, $p = 0,045$), inisiasi menyusui dini (IMD) ($OR = 0,76$, $p = 0,024$), konsumsi tablet tambah darah (TTD) ($OR = 5,9$, $p = 0,02$), ASI eksklusif ($OR = 0,6$, $p = 0,015$), MP-ASI ($OR = 0,678$, $p = 0,039$), dan imunisasi ($OR = 1,6$, $p = 0,016$) merupakan faktor-faktor yang mempengaruhi kejadian stunting secara signifikan.

[Amalina et al. \(2026\)](#) dalam skripsinya yang berjudul *Optimisasi Analisis Regresi Logistik LASSO Menggunakan Bayesian Information Criterion (BIC): Studi Kasus Faktor Penyebab Perceraian di Indonesia Tahun 2024* menerapkan regresi logistik LASSO dengan optimasi BIC untuk mengidentifikasi faktor-faktor penyebab perceraian di Indonesia. Penelitian ini menggunakan tiga belas variabel prediktor dan menemukan bahwa model regresi logistik awal mengalami multikolinearitas, ditandai dengan nilai VIF yang melebihi 10 pada beberapa variabel. Hasil analisis LASSO-BIC menunjukkan bahwa nilai BIC minimum sebesar 26,770 tercapai pada lambda optimal sebesar 0,003914. Model akhir mempertahankan enam variabel prediktor, yaitu judi, ditinggalkan salah satu pihak, salah satu pihak dipenjara, poligami, cacat badan atau penyakit, dan perselisihan. Seluruh koefisien variabel tersebut bernilai positif. Variabel cacat badan atau penyakit memiliki koefisien terbesar sebesar 5,764, diikuti oleh perselisihan sebesar 5,198, yang menunjukkan bahwa kedua faktor ini paling dominan dalam menentukan tingkat perceraian tinggi di suatu provinsi. Penelitian ini juga membandingkan model LASSO-BIC dengan regresi logistik konvensional (MLE). Model MLE menghasilkan 14 parameter dengan BIC sebesar 66,055, sedangkan model LASSO-BIC hanya menghasilkan 7

parameter dengan BIC sebesar 26,770. Hasil ini membuktikan bahwa LASSO-BIC menghasilkan model yang lebih sederhana dan lebih baik dalam menyeimbangkan kesesuaian model dengan kompleksitas struktur variabel.

Wijayakanthi et al. (2026) dalam skripsinya yang berjudul *Optimasi Regresi Logistik LASSO dengan Kriteria Bayesian Information Criterion (BIC): Studi Kasus Klasifikasi Kemiskinan di Indonesia Tahun 2024* menerapkan regresi logistik LASSO dengan optimasi BIC untuk mengidentifikasi faktor penentu klasifikasi kemiskinan pada 38 provinsi di Indonesia. Penelitian ini menggunakan data dari Badan Pusat Statistik dengan sembilan variabel independen yang mencakup faktor sosial, ekonomi, kesehatan, pendidikan, dan ketenagakerjaan. Hasil analisis menunjukkan bahwa nilai BIC minimum sebesar 41,695 tercapai pada lambda optimal 0,01737. Model akhir mempertahankan tiga variabel prediktor, yaitu Umur Harapan Hidup (UHH), Tingkat Partisipasi Angkatan Kerja (TPAK), dan Produk Domestik Regional Bruto (PDRB). Variabel UHH dan PDRB berpengaruh negatif terhadap probabilitas klasifikasi kemiskinan di atas rata-rata, dengan koefisien masing-masing sebesar -1,591 dan -1,224. Sebaliknya, TPAK berpengaruh positif dengan koefisien sebesar 1,100. Penelitian ini juga membandingkan performa model LASSO dengan regresi logistik konvensional (MLE) hasil eliminasi backward stepwise. Model MLE tercatat memiliki BIC lebih rendah (40,509) dibandingkan model LASSO (41,695), namun model LASSO tetap unggul dalam menangani multikolinearitas dan menghasilkan koefisien yang lebih stabil untuk tujuan prediksi.

Berdasarkan tinjauan pustaka di atas, terdapat beberapa kesenjangan yang belum sepenuhnya dieksplorasi dalam penelitian-penelitian sebelumnya. Pertama, (Safitri et al., 2022) dan (Mataheru & Utari, 2026) telah menggunakan regresi logistik biner untuk mengidentifikasi faktor risiko stunting, namun kedua penelitian

tersebut tidak mengaplikasikan metode regularisasi seperti LASSO yang mampu menangani overfitting pada data dengan banyak variabel prediktor yang berkorelasi. Kedua, (Mkungudza et al., 2024) telah membuktikan bahwa LASSO efektif untuk seleksi variabel dalam prediksi stunting di Malawi, namun penelitian tersebut tidak mengoptimasi parameter penalti λ dengan AIC melainkan dengan cross-validation. Ketiga, (Sari et al., 2026) dan (Arofah & Sofro, 2022) telah mendemonstrasikan penggunaan AIC untuk pemilihan model dan seleksi variabel, namun belum ada yang mengkombinasikan AIC dengan LASSO pada data stunting. Keempat, (Mochtar et al., 2024) telah menganalisis intervensi gizi spesifik, namun dengan ukuran sampel yang relatif kecil (40 responden). Kelima, (Rusdy et al., 2025) telah berhasil mengkombinasikan GWLR dengan LASSO, namun pendekatan serupa belum diterapkan pada pemodelan faktor risiko stunting. Keenam, (Amalina et al., 2026) dan (Wijayakanthi et al., 2026) telah menerapkan regresi logistik LASSO dengan optimasi parameter penalti menggunakan BIC pada kasus perceraian dan klasifikasi kemiskinan di Indonesia, namun kedua penelitian tersebut menggunakan BIC dan bukan AIC sebagai kriteria pemilihan model, serta belum diterapkan pada konteks data kesehatan seperti stunting. Ketujuh, belum ada penelitian yang secara khusus mengaplikasikan regresi logistik LASSO dengan optimasi AIC pada data Survei Status Gizi Indonesia (SSGI) skala nasional. Kesenjangan inilah yang menjadi dasar dan kontribusi utama penelitian ini.

1.7. Sistematika Penulisan

Sistematika pada penelitian ini yaitu:

BAB 1 PENDAHULUAN

Pada bab ini berisi tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, sistematika penulisan, dan tinjauan pustaka.

BAB II DASAR TEORI

Pada bab ini menjelaskan terkait landasan teori yang menjadi acuan, berisi tentang konsep, hipotesis dan rancangan penelitian.

BAB III METODE PENELITIAN

Pada bab ini membahas terkait metode penelitian yang akan digunakan dalam pemecahan masalah yang berisi tentang pendekatan penelitian, sumber data, metode pengolahan data dan diagram analisis data.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini akan dilakukan proses analisis data yang kemudian hasilnya diinterpretasikan secara terperinci.

BAB V PENUTUP

Pada bab ini berisi kesimpulan dan saran dari penelitian yang dilakukan oleh penulis.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil penelitian mengenai optimisasi regresi logistik LASSO dengan kriteria *Akaike Information Criterion* (AIC) untuk klasifikasi faktor risiko stunting berdasarkan data *Survei Status Gizi Indonesia* (SSGI) Tahun 2024, diperoleh kesimpulan sebagai berikut.

1. Regresi logistik LASSO dapat digunakan untuk membangun model klasifikasi pada data stunting yang memiliki banyak variabel prediktor. Metode ini melakukan seleksi variabel melalui pemberian penalti terhadap koefisien regresi sehingga koefisien yang kurang berkontribusi akan menyusut menjadi nol. Dengan demikian, model yang dihasilkan menjadi lebih sederhana dan lebih mudah diinterpretasikan dibandingkan model regresi logistik penuh.
2. Optimisasi model regresi logistik LASSO menggunakan *Akaike Information Criterion* (AIC) berhasil memperoleh nilai parameter regularisasi (λ) optimum yang menghasilkan model terbaik berdasarkan nilai AIC minimum. Penggunaan AIC mampu menyeimbangkan antara tingkat kecocokan model dan kompleksitas model sehingga hanya variabel yang memberikan kontribusi penting yang dipertahankan dalam model akhir.
3. Penerapan regresi logistik LASSO dengan optimisasi AIC pada data SSGI Tahun 2024 berhasil mengatasi permasalahan multikolinearitas yang muncul

pada model regresi logistik penuh. Dari 19 variabel prediktor awal, model akhir mempertahankan tujuh variabel, yaitu *Jarak Kehamilan Risiko*, *Berat Badan Lahir Rendah (BBLR)*, *ASI Eksklusif*, *Minimum Dietary Diversity (MDD)*, *Imunisasi Lengkap*, *ISPA*, dan *Pendidikan Ibu Tinggi*. Berdasarkan nilai *odds ratio*, variabel *Jarak Kehamilan Risiko* dan *BBLR* merupakan faktor yang meningkatkan peluang suatu provinsi berada pada kategori prevalensi stunting tinggi. Sebaliknya, variabel *ASI Eksklusif*, *MDD*, *Imunisasi Lengkap*, dan *Pendidikan Ibu Tinggi* berperan sebagai faktor protektif terhadap stunting. Variabel *ISPA* tetap dipertahankan dalam model, namun interpretasinya perlu dilakukan secara hati-hati karena penelitian menggunakan data agregat tingkat provinsi.

Evaluasi model menghasilkan nilai *accuracy*, *sensitivity*, *specificity*, dan *precision* sebesar 100%, serta nilai *Area Under the Curve (AUC)* sebesar 1,00. Hasil tersebut menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik terhadap data penelitian. Namun demikian, hasil evaluasi tersebut perlu diinterpretasikan secara hati-hati karena proses pembentukan dan evaluasi model dilakukan menggunakan data yang sama serta jumlah observasi penelitian relatif terbatas.

Secara keseluruhan, penelitian ini menunjukkan bahwa regresi logistik LASSO dengan optimisasi *Akaike Information Criterion (AIC)* mampu menghasilkan model klasifikasi faktor risiko stunting yang lebih sederhana, lebih stabil, dan tetap memiliki kemampuan klasifikasi yang sangat baik. Model yang dihasilkan dapat menjadi salah satu pendekatan statistik dalam mengidentifikasi faktor-faktor dominan yang berkaitan dengan prevalensi stunting di Indonesia.

5.2. Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat diberikan adalah sebagai berikut.

1. Penelitian selanjutnya disarankan menggunakan jumlah observasi yang lebih besar, misalnya data pada tingkat individu atau rumah tangga, sehingga model yang diperoleh dapat menggambarkan hubungan antarvariabel secara lebih akurat dan mengurangi potensi *ecological fallacy* akibat penggunaan data agregat tingkat provinsi.
2. Penelitian selanjutnya disarankan melakukan validasi model menggunakan data latih dan data uji atau metode *cross-validation* sehingga kemampuan generalisasi model terhadap data baru dapat dievaluasi secara lebih objektif.
3. Penelitian selanjutnya dapat membandingkan optimisasi parameter regularisasi menggunakan kriteria pemilihan model lain, seperti *Bayesian Information Criterion* (BIC) atau *Corrected Akaike Information Criterion* (AICc), untuk mengetahui konsistensi variabel yang terpilih dan memperoleh model yang lebih optimal.
4. Bagi pemerintah dan pemangku kebijakan, hasil penelitian ini menunjukkan bahwa upaya penanggulangan stunting perlu mempertimbangkan faktor-faktor yang berhubungan dengan peluang prevalensi stunting tinggi. Jarak kehamilan berisiko dan *Berat Badan Lahir Rendah* (BBLR) perlu menjadi perhatian utama karena menunjukkan hubungan positif dengan peluang prevalensi stunting tinggi dalam model. Sementara itu, ISPA, ASI eksklusif, *Minimum Dietary Diversity* (MDD), imunisasi lengkap, dan pendidikan ibu

tinggi dapat menjadi faktor tambahan yang dipertimbangkan dalam penyusunan strategi percepatan penurunan prevalensi stunting di Indonesia.



DAFTAR PUSTAKA

- Agresti, A. (2013). *Categorical Data Analysis*. John Wiley & Sons, Hoboken, NJ, 3rd edition, ISBN: [978-0470463635](#).
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, DOI: [10.1109/TAC.1974.1100705](#).
- Al-khateeb, S. F. M. (2019). Tuning parameter selection in penalized logistic regression with application in cancer. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 18(36):11–22.
- Amalina, M. N. et al. (2026). *OPTIMISASI ANALISIS REGRESI LOGISTIK LASSO MENGGUNAKAN BAYESIAN INFORMATION CRITERION (BIC)(STUDI KASUS: FAKTOR PENYEBAB PERCERAIAN DI INDONESIA TAHUN 2024)*. PhD thesis, UIN SUNAN KALIJAGA YOGYAKARTA.
- Arofah, F. & Sofro, A. (2022). Penerapan regresi logistik multinomial untuk analisis model tingkat depresi pada lansia. *MATHunesa: Jurnal Ilmiah Matematika*, 10(1):84–93.
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons.
- Burnham, K. P. & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, New York, 2nd edition, DOI: [10.1007/b97636](#).

- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22, DOI: [10.18637/jss.v033.i01](https://doi.org/10.18637/jss.v033.i01).
- Greene, W. H. (2020). *Econometric Analysis*. Pearson, Boston, MA, 8th edition.
- Gujarati, D. N. & Porter, D. C. (2009). *Basic Econometrics*. McGraw-Hill, New York, 5th edition.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, 2nd edition, ISBN: [978-0-387-84857-0](https://doi.org/10.1007/978-0-387-84857-0), DOI: [10.1007/978-0-387-84858-7](https://doi.org/10.1007/978-0-387-84858-7).
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. John Wiley & Sons, Hoboken, NJ, 3rd edition, DOI: [10.1002/9781118548387](https://doi.org/10.1002/9781118548387).
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: with Applications in R*. Springer, New York, 2nd edition, ISBN: [978-1-0716-1418-1](https://doi.org/10.1007/978-1-0716-1418-1), DOI: [10.1007/978-1-0716-1418-1](https://doi.org/10.1007/978-1-0716-1418-1).
- Kemenkes (2024). Survei status gizi indonesia (ssgi) 2024: Ssgi dalam angka. Technical report, Badan Kebijakan Pembangunan Kesehatan, Kementerian Kesehatan RI, Jakarta.
- Mataheru, J. S. & Utari, D. M. (2026). Faktor risiko yang berhubungan dengan kejadian stunting pada balita di sulawesi barat: Analisis data ssgi 2024. *Malahayati Nursing Journal*, 8(6):12–27.
- McCullagh, P. & Nelder, J. A. (1989). *Generalized Linear Models*. Chapman and Hall/CRC, London, 2nd edition, DOI: [10.1201/9780203753736](https://doi.org/10.1201/9780203753736).

- Mkungudza, J., Twabi, H. S., & Manda, S. O. (2024). Development of a diagnostic predictive model for determining child stunting in malawi: a comparative analysis of variable selection approaches. *BMC Medical Research Methodology*, 24(1):175.
- Mochtar, N., Ali, N. M., Mursaha, A., et al. (2024). Analisis regresi logistik biner pada karakteristik demografi sosial ekonomi dan intervensi gizi spesifik balita stunting di wilayah kerja puskesmas jambula kota ternate. *Hospital Majapahit (JURNAL ILMIAH KESEHATAN POLITEKNIK KESEHATAN MOJOKERTO)*, 16(2):148–166.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons, 5th edition.
- Nelder, J. A. & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, 135(3):370–384, DOI: [10.2307/2344614](https://doi.org/10.2307/2344614).
- R Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>.
- Rusdy, D. A. P. S., Thamrin, S. A., & Islamiyati, A. (2025). Determining factors that influence unmet need for family planning using geographically weighted logistic regression with lasso. *Jurnal Matematika, Statistika dan Komputasi*, 21(2):583–600.
- Safitri, E., Basriati, S., & Mulyani, S. (2022). Faktor-faktor yang mempengaruhi kejadian stunting terhadap balita menggunakan analisis regresi logistik. *Zeta-Math Journal*, 7(2):47–52.

- Sari, D. N., Samodra, G., & Hadmoko, D. S. (2026). Pemodelan spasial kerawanan longsor di kabupaten temanggung menggunakan regresi logistik dengan seleksi variabel stepwise aic. *JPIG (Jurnal Pendidikan dan Ilmu Geografi)*, 11(1):32–43.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, DOI: [10.1111/j.2517-6161.1996.tb02080.x](https://doi.org/10.1111/j.2517-6161.1996.tb02080.x).
- WHO (2014). Stunting in a nutshell. <https://www.who.int/news-room/questions-and-answers/item/stunting>. Diakses 2 April 2025.
- Wijayakanthi, F. A. et al. (2026). *OPTIMASI REGRESI LOGISTIK LASSO DENGAN KRITERIA BAYESIAN INFORMATION CRITERION (BIC)(STUDI KASUS: KLASIFIKASI KEMISKINAN DI INDONESIA TAHUN 2024)*. PhD thesis, UIN SUNAN KALIJAGA YOGYAKARTA.
- Zou, H., Hastie, T., & Tibshirani, R. (2007). On the degrees of freedom of the lasso. *The Annals of Statistics*, 35(5):2173–2192.