

LAPORAN SKRIPSI

KETEPATAN KLASIFIKASI STATUS *STUNTING* PADA BALITA MENGGUNAKAN METODE REGRESI LOGISTIK BINER DAN *CLASSIFICATION AND REGRESSION TREE* (CART) BERDASARKAN NILAI *F1-SCORE*

(Studi Kasus : Kasus *Stunting* di Papua Tengah Tahun 2023)



NOORKHALIZA MAULIDYANI SAFIRA

NIM. 21106010072

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2025

SKRIPSI

KETEPATAN KLASIFIKASI STATUS *STUNTING* PADA BALITA MENGGUNAKAN METODE REGRESI LOGISTIK BINER DAN *CLASSIFICATION AND REGRESSION TREE* (*CART*) BERDASARKAN NILAI *F1-SCORE*

(Studi Kasus : Kasus *Stunting* di Papua Tengah Tahun 2023)

Diajukan untuk memenuhi salah satu syarat memperoleh derajat

Sarjana Matematika



NOORKHALIZA MAULIDYANI SAFIRA

NIM. 21106010072

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2025



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 : Noorkhaliza Maulidyani Safira

NIM : 21106010072

Judul Skripsi : Ketepatan Klasifikasi Status *Stunting* Pada Balita Menggunakan Metode Regresi Logistik Biner dan *Classification and Regression Tree (CART)* Berdasarkan Nilai *F1-score* (Studi Kasus: Kasus *Stunting* di Papua Tengah Tahun 2023)

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, 07 Juli 2025

Dosen Pembimbing Skripsi I

Pipit Pratiwi Rahayu, S.Si., M.Sc.

NIP. 19861208 201503 2 006

Dosen Pembimbing Skripsi II

Muhamad Rashif Hilmi, S.Si., M.Sc.

NIP. 19920309 202012 1 001



PENGESAHAN TUGAS AKHIR

Nomor : B-1619/Un.02/DST/PP.00.9/08/2025

Tugas Akhir dengan judul : Ketepatan Klasifikasi Status Stunting Pada Balita Menggunakan Metode Regresi Logistik Biner dan Classification and Regression Tree (CART) Berdasarkan Nilai F1-score (Studi Kasus: Kasus Stunting di Papua Tengah Tahun 2023)

yang dipersiapkan dan disusun oleh:

Nama : NOORKHALIZA MAULIDYANI SAFIRA
Nomor Induk Mahasiswa : 21106010072
Telah diujikan pada : Rabu, 16 Juli 2025
Nilai ujian Tugas Akhir : A

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

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Pipit Pratiwi Rahayu, S.Si., M.Sc.
SIGNED

Valid ID: 68898824314ce



Penguji I

Muhamad Rashif Hilmi, S.Si., M.Sc.
SIGNED

Valid ID: 6891b7a7166c2



Penguji II

Dr. Epha Diana Supandi, S.Si., M.Sc.
SIGNED

Valid ID: 6886e8c85de8f



Yogyakarta, 16 Juli 2025
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

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

Valid ID: 6892ff19d1b4a

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Noorkhaliza Maulidyani Safira

NIM : 21106010072

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, 07 Juli 2025



Noorkhaliza Maulidyani Safira

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur dan kerendahan hati, skripsi ini saya persembahkan kepada kedua orang tua tercinta, para kakak kandung terkeren, keluarga besar dari ayah dan ibu, serta sahabat dan teman-teman seperjuangan. Terima kasih atas doa dan dukungannya, sehingga karya yang penuh perjuangan ini bisa terwujud.



HALAMAN MOTTO

Berbuatlah baik bukan karena mereka baik, tetapi karena memang kamu baik.

-

Whatever happens, be true to your heart.

-

“Beri hati pada setiap kerja kerasmu, karya-karyamu”

- Mahakarya by Tulus -

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

KATA PENGANTAR

Alhamdulillah, puji syukur kepada Allah SWT yang telah melimpahkan Rahmat dan karunia-Nya berupa keimanan, kekuatan, kesabaran, dan kelancaran sehingga penulis dapat menyelesaikan skripsi yang berjudul “Ketepatan Klasifikasi Status *Stunting* pada Balita Menggunakan Metode Regresi Logistik Biner dan *Classification and Regression Tree* (CART) Berdasarkan Nilai *F1-score* (Studi Kasus : Kasus *Stunting* di Papua Tengah Tahun 2023)”. Sholawat serta salam senantiasa tercurahkan kepada Nabi Muhammad SAW yang kita nantikan syafa’atnya di yaumul qiyamah nanti.

Skripsi ini ditulis sebagai syarat dalam menyelesaikan studi S-1 Matematika di Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta. Penulis menyadari bahwa selama penyusunan skripsi ini telah banyak mendapatkan doa, bantuan, bimbingan serta dukungan dari berbagai pihak. Untuk itu pada kesempatan ini, penulis bermaksud ingin menyampaikan ucapan terima kasih kepada:

1. Allah SWT. yang senantiasa melimpahkan segala rahmat, hidayah, serta kekuatan, sehingga penulis dapat menyelesaikan penyusunan skripsi ini dengan baik.
2. Prof. Noorhaidi Hasan, M.A., M.Phil., Ph.D. selaku Rektor UIN Sunan Kalijaga Yogyakarta.
3. Prof. Dr. Khurul Wardati, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
4. Dr. Epha Diana Supandi, S.Si., M.Sc., selaku Ketua Program Studi Matematika UIN Sunan Kalijaga Yogyakarta.
5. Pipit Pratiwi Rahayu, S.Si., M.Sc., selaku Dosen Penasihat Akademik dan Dosen Pembimbing Skripsi I yang telah memberikan arahan, masukan dan nasihat kepada penulis selama dibangku perkuliahan.
6. Muhamad Rashif Hilmi, S.Si., M.Sc., selaku Dosen Pembimbing Skripsi II yang telah berjasa dan selalu membimbing penulis dalam penyusunan skripsi ini dengan sangat baik.

7. Seluruh dosen Matematika dan staf Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta yang selalu memberikan ilmu dan pelayanan yang baik selama perkuliahan.
8. Kedua orang tua tercinta, Ayah Drs. H. Khairuddin Hasnawi, M.M. dan Ibu Dra. Hj. Muzdalifah yang senantiasa memberikan kasih sayang, doa, dan dukungan tiada henti sepanjang hidup penulis hingga saat ini. Terima kasih atas segala perjuangan ayah dan ibu dalam mengusahakan serta memberikan yang terbaik bagi penulis. Semoga penulis dapat membuat ayah dan ibu turut menyaksikan dan merasakan kebahagiaan atas hasil dari setiap doa dan usaha yang telah diberikan.
9. Keempat kakak kandung penulis – cak Rikha, abang Haris, woh Santi, dan cek Rani– yang selalu memberikan semangat, dorongan, dan motivasi selama masa perkuliahan hingga penyusunan skripsi ini. Terima kasih telah menjadi tempat pulang, berbagi cerita, dan sumber kekuatan di setiap langkah perjuangan ini. Terima kasih pula atas segala saran yang diberikan saat adik bungsu kalian ini dilanda kebingungan.
10. Ketiga “kurcaci” yang bermula dari grup ‘*hiling*’ – Ake, Visca, dan Fitri. Terima kasih atas kebersamaannya di perantauan ini dengan segala warna yang telah kita ciptakan. Terima kasih sudah saling mengusahakan dan merayakan satu sama lain. As you know guys, dipertemukan dengan kalian di perkuliahan ini adalah salah satu rezeki terbaik yang Allah berikan. Penulis selalu berharap, pertemuan ini tidak berhenti sampai di sini saja, tetapi terus berlanjut hingga kita tua nanti.
11. Teman-teman Matematika 2021, terutama pada konsentrasi Statistika yang sudah kebersamai selama masa perkuliahan. Terima kasih sudah saling membantu dan menyemangati di setiap semesternya. Semoga di setiap langkah kedepannya selalu dilancarkan.
12. Teman-teman seperbimbingan ‘Skripsi Matematika UIN Suka Ceria’. Terima kasih sudah saling menyemangati dan saling bertukar informasi dengan baik.
13. Partner PDD selama KKN, Ineke, yang meskipun jauh di Kediri, tetapi kita tetap saling berbagi cerita, semangat, dan dukungan hingga saat ini. Terima

kasih atas segala kerja sama yang pernah kita jalani bersama selama masa KKN. Semoga hubungan baik ini terus terjaga, dan kita bisa saling mendukung dalam setiap langkah kehidupan ke depan.

14. Teman-teman HMPS Matematika Tahun 2022 dan 2023, terutama di departemen Minat Bakat dan Cyber Space. Terima kasih telah memberikan banyak ilmu serta pengalaman yang membuat penulis semakin berkembang terkhusus dalam dunia desain. Semoga ilmu yang diperoleh dapat bermanfaat untuk penulis kedepannya.
15. Diri sendiri, Noorkhaliza Maulidyani Safira. Apresiasi sebesar-besarnya karena telah berjuang dan menyelesaikan apa yang telah dimulai. Terima kasih sudah terus berusaha memberikan yang terbaik dan menyediakan ruang untuk ikhlas atas segala hal yang terjadi.
16. Semua pihak yang tidak dapat disebutkan satu per satu yang telah membantu penulis dalam menyelesaikan skripsi ini. Terima kasih atas segala doa dan bantuannya, semoga Allah SWT membalas kebaikan kalian.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, sehingga penulis mengharap kritik dan saran dari semua pihak guna kebaikan dari skripsi ini.

Yogyakarta, 15 Mei 2025

Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
SURAT PERNYATAAN KEASLIAN	v
HALAMAN PERSEMBAHAN	vi
HALAMAN MOTTO	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvi
DAFTAR SIMBOL	xvii
INTISARI.....	xix
ABSTRACT.....	xx
BAB 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	6
1.6 Tinjauan Pustaka	6
1.7 Sistematika Penulisan	10
BAB II LANDASAN TEORI	12
2.1 Matriks	12
2.1.1 Jenis Matriks	12
2.1.2 Operasi Matriks.....	13
2.1.3 Transpose Matriks	15
2.1.4 Invers Matriks	15
2.2 Teori Probabilitas	15

2.3	Distribusi Probabilitas.....	16
2.3.1	Distribusi Bernoulli.....	17
2.3.2	Distribusi Binomial.....	18
2.4	<i>Data Mining</i>	19
2.5	Klasifikasi.....	21
2.6	Analisis Regresi.....	21
2.6.1	Analisis Regresi Linier Berganda.....	22
2.6.2	Analisis Regresi Logistik.....	23
2.7	<i>Decission Tree</i>	24
2.8	<i>Stunting</i>	26
2.8.1	Penentuan Status <i>Stunting</i> dengan <i>Z-score</i>	26
2.8.2	Faktor Penyebab <i>Stunting</i>	27
BAB III METODE PENELITIAN.....		31
3.1	Jenis Penelitian.....	31
3.2	Jenis dan Sumber Data.....	31
3.3	Variabel Penelitian.....	31
3.4	Alat Pengolahan Data.....	32
3.5	Metode Analisis Data.....	33
3.6	<i>Flowchart</i> Penelitian.....	34
BAB IV PEMBAHASAN.....		35
4.1	Regresi Logistik Biner.....	35
4.1.1	Model Regresi Logistik Biner.....	35
4.1.2	Uji Multikolinieritas.....	36
4.1.3	Estimasi Parameter.....	37
4.1.4	Uji Signifikansi Parameter.....	42
4.1.5	Interpretasi Koefisien Parameter.....	44
4.2	<i>Classification and Regression Tree (CART)</i>	45
4.2.1	Pembentukan Pohon Klasifikasi.....	47
4.2.2	Pemangkasan Pohon Klasifikasi.....	50
4.2.3	Penentuan Pohon Klasifikasi Optimal.....	51
4.3	Perhitungan Ketepatan Klasifikasi (<i>F1-score</i>).....	52
BAB V STUDI KASUS.....		54

5.1	<i>Data Preprocessing</i>	54
5.1.1	<i>Data Cleaning</i>	54
5.1.2	<i>Data Transformation</i>	55
5.1.3	Penanganan <i>Outlier</i>	57
5.2	Eksplorasi Data	58
5.3	Pembagian Data	61
5.4	Analisis Regresi Logistik Biner	62
5.4.1	Uji Multikolinieritas	62
5.4.2	Estimasi Parameter	63
5.4.3	Uji Signifikansi Parameter	65
5.4.4	Model Regresi Logistik Biner	68
5.4.5	Interpretasi Koefisien	69
5.4.6	Ketepatan Klasifikasi Regresi Logistik Biner	70
5.5	Analisis <i>Classification and Regression Tree</i> (CART)	71
5.5.1	Pembentukan Pohon Klasifikasi Maksimal	71
5.5.2	Pemangkasan Pohon Klasifikasi Maksimal	74
5.5.3	Pohon Klasifikasi Optimal	75
5.5.4	Ketepatan Klasifikasi CART	77
5.6	Perbandingan Nilai Ketepatan Klasifikasi (<i>F1-score</i>)	78
5.7	Klasifikasi Data Baru dengan Metode Terbaik	79
BAB VI PENUTUP		81
6.1	Kesimpulan	81
6.2	Saran	82
DAFTAR PUSTAKA		83
LAMPIRAN		86

DAFTAR TABEL

Tabel 1.1 Tinjauan pustaka.....	8
Tabel 2.1 Kategori status <i>stunting</i>	27
Tabel 3.1 Variabel penelitian.....	31
Tabel 4.1 <i>Confusion matrix</i>	52
Tabel 5.1 Deskriptif data usia, berat badan, dan tinggi badan	61
Tabel 5.2 Nilai VIF.....	63
Tabel 5.3 Estimasi parameter	63
Tabel 5.4 Uji Wald pertama.....	66
Tabel 5.5 Uji Wald kedua	67
Tabel 5.6 Uji signifikansi simultan	68
Tabel 5.7 Koefisien regresi	69
Tabel 5.8 Nilai <i>odds ratio</i>	69
Tabel 5.9 Ketepatan klasifikasi regresi logistik biner	70
Tabel 5.10 Perhitungan kemungkinan pemilah.....	71
Tabel 5.11 Skor variabel terpenting pohon klasifikasi maksimal	72
Tabel 5.12 Kondisi simpul pohon klasifikasi maksimal	74
Tabel 5.13 Tabel simpul terminal dan penandaan label kelas	74
Tabel 5.14 Nilai <i>complexity parameter</i>	75
Tabel 5.15 Ketepatan klasifikasi CART.....	78
Tabel 5.16 Perbandingan nilai <i>F1-score</i>	78
Tabel 5.17 Data baru	79
Tabel 5.18 Hasil klasifikasi data baru	80

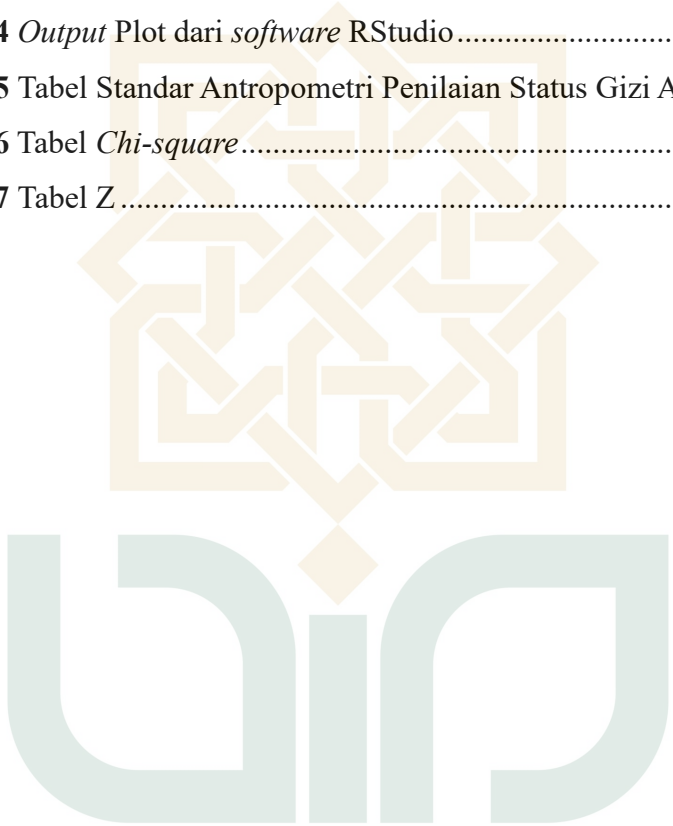
DAFTAR GAMBAR

Gambar 1.1 Persentase <i>stunting</i> di Indonesia tahun 2013-2023	2
Gambar 1.2 Prevalensi <i>stunting</i> menurut provinsi Indonesia tahun 2022-2023	2
Gambar 2.1 Fungsi regresi logistik	24
Gambar 2.2 Model <i>decision tree</i>	25
Gambar 3.1 <i>Flowchart</i> penelitian	34
Gambar 4.1 Skema pohon metode CART	47
Gambar 5.1 <i>Output missing value</i>	54
Gambar 5.2 Input data pada <i>software</i> WHO Anthro	56
Gambar 5.3 Output nilai <i>z-score</i> pada <i>software</i> WHO Anthro	56
Gambar 5.4 <i>Boxplot outlier</i> pada variabel numerik	57
Gambar 5.5 Persentase balita terkait status <i>stunting</i>	59
Gambar 5.6 Persebaran <i>stunting</i> tiap kabupaten di Papua Tengah	59
Gambar 5.7 Persebaran <i>stunting</i> berdasarkan daerah tempat tinggal	60
Gambar 5.8 Persebaran <i>stunting</i> berdasarkan jenis kelamin	60
Gambar 5.9 Pohon klasifikasi maksimal	73
Gambar 5.10 Pohon klasifikasi optimal	76

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR LAMPIRAN

Lampiran 1 Data dan <i>Script</i> R.....	86
Lampiran 2 Keterangan Variabel Dummy.....	86
Lampiran 3 <i>Input</i> dan <i>Output</i> Penggunaan <i>software</i> WHO Anthro	87
Lampiran 4 <i>Output</i> Plot dari <i>software</i> RStudio.....	88
Lampiran 5 Tabel Standar Antropometri Penilaian Status Gizi Anak.....	90
Lampiran 6 Tabel <i>Chi-square</i>	91
Lampiran 7 Tabel Z.....	92



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR SIMBOL

Y	= variabel respon
X	= variabel prediktor
β_0	= konstanta
β_p	= parameter untuk variabel prediktor ke- i , $i = 1, 2, \dots, p$
$\hat{\beta}_i$	= nilai estimasi dari parameter β_i , $i = 1, 2, \dots, p$
ε	= galat atau <i>error</i>
p	= banyaknya variabel prediktor
n	= total pengamatan
χ^2	= distribusi <i>chi-square</i>
$P(x)$	= probabilitas suatu hasil tertentu dalam distribusi probabilitas
$\pi(x_i)$	= probabilitas sukses kejadian dengan nilai probabilitas $0 \leq \pi(x_i) \leq 1$
$g(x)$	= fungsi logit
$L(\beta)$	= fungsi <i>likelihood</i>
$l(\beta)$	= fungsi log <i>likelihood</i>
ψ	= <i>odds ratio</i>
$SE(\hat{\beta}_i)$	= estimasi alat baku untuk koefisien β_i , $i = 1, 2, \dots, p$
L_0	= <i>likelihood</i> model tanpa variabel prediktor
L_p	= <i>likelihood</i> model dengan p variabel prediktor
$P(j t)$	= proporsi kelas j pada simpul t
$n_j(t)$	= jumlah pengamatan kelas j pada simpul t
$n(t)$	= jumlah pengamatan pada simpul t
$i(t)$	= impuritas pada simpul t
$\Delta i(s, t)$	= penurunan impuritas akibat pemisahan s di simpul t
$i(t_R)$	= impuritas simpul kanan setelah <i>split</i>
$i(t_L)$	= impuritas simpul kiri setelah <i>split</i>
P_R	= proporsi data yang masuk ke simpul kanan dari simpul t
P_L	= proporsi data yang masuk ke simpul kiri dari simpul t
$N_j(t)$	= jumlah pengamatan kelas j pada simpul t

$N(t)$ = jumlah pengamatan pada simpul t
 $g_m(t)$ = fungsi *impurity* pada simpul t
 $R(t)$ = kesalahan pengklasifikasian pada simpul t
 $R(T_k)$ = kesalahan pengklasifikasian pada pohon T_k
 $R(t_{m_k})$ = kesalahan pengklasifikasian pada simpul t ke- m di pohon ke- k
 T_k = *subtree* ke- k , $k = 1, 2, 3, \dots, n$
 $|T_k|$ = jumlah simpul terminal pada *subtree* T_k
 V = jumlah lipatan dalam *V-fold cross validation*
 L_v = bagian data ke- v yang dijadikan sebagai data uji pada lipatan ke- v
 N_v = jumlah pengamatan dalam L_v
 $T^{(v)}$ = pohon keputusan yang dibentuk dari data $L - L_v$
 $d^{(v)}$ = hasil klasifikasi pada data uji L_v oleh pohon $T^{(v)}$
 $X(d^{(v)})$ = hasil pengklasifikasian
 $R(T_t^{(v)})$ = estimasi kesalahan rata-rata dari seluruh lipatan

INTISARI

KETEPATAN KLASIFIKASI STATUS *STUNTING* PADA BALITA MENGUNAKAN METODE REGRESI LOGISTIK BINER DAN *CLASSIFICATION AND REGRESSION TREE (CART)* BERDASARKAN NILAI *F1-SCORE*

(Studi Kasus : Kasus *Stunting* di Papua Tengah tahun 2023)

Oleh
NOORKHALIZA MAULIDYANI SAFIRA
NIM. 21106010072

Stunting merupakan masalah gizi kronis pada balita yang berdampak jangka panjang terhadap pertumbuhan fisik, perkembangan kognitif, serta produktivitas di masa depan. Papua Tengah tercatat sebagai provinsi dengan prevalensi *stunting* tertinggi di Indonesia pada tahun 2023, yaitu sebesar 39,4%. Penelitian ini bertujuan untuk mengidentifikasi variabel yang berpengaruh terhadap kejadian *stunting* dan membandingkan ketepatan klasifikasi menggunakan metode Regresi Logistik Biner dan *Classification and Regression Tree (CART)* berdasarkan nilai *F1-score*. Data yang digunakan bersumber dari Survei Kesehatan Indonesia (SKI) tahun 2023 yang diperoleh dari Pusat Data dan Teknologi Informasi (PUSDATIN) Kemenkes RI, dengan total sampel sebanyak 401 balita dari enam kabupaten di Papua Tengah. Pengolahan data dan analisis dilakukan menggunakan perangkat lunak RStudio. Status *stunting* ditentukan berdasarkan nilai *z-score* indeks tinggi badan menurut umur (TB/U) yang dihitung menggunakan perangkat lunak WHO Anthro. Hasil analisis menunjukkan bahwa dari 11 variabel prediktor, hanya variabel usia dan tinggi badan yang signifikan terhadap *stunting*. Nilai *F1-score* untuk metode Regresi Logistik Biner dan metode CART berturut-turut sebesar 95.73% dan 90.38%. Dengan demikian, metode Regresi Logistik Biner lebih optimal dalam mengidentifikasi faktor-faktor yang berpengaruh terhadap *stunting* pada balita di Papua Tengah tahun 2023.

Kata kunci: *Stunting*, Regresi Logistik Biner, CART, Klasifikasi, Papua Tengah.

ABSTRACT

ACCURACY OF STUNTING STATUS CLASSIFICATION IN TODDLERS USING BINARY LOGISTIC REGRESSION AND CLASSIFICATION AND REGRESSION TREE (CART) METHODS BASED ON F1-SCORE VALUES

(Case Study: Stunting Cases in Central Papua in 2023)

by

NOORKHALIZA MAULIDYANI SAFIRA
NIM. 21106010072

Stunting is a chronic nutritional problem in toddlers that has long-term effects on physical growth, cognitive development, and future productivity. Central Papua is recorded as the province with the highest stunting prevalence in Indonesia in 2023, at 39.4%. This study aims to identify variables influencing stunting incidence and compare the accuracy of classification using Binary Logistic Regression and Classification and Regression Tree (CART) methods based on F1-score values. The data used were sourced from the 2023 Indonesian Health Survey obtained from the Center for Data and Information Technology of the Indonesian Ministry of Health, with a total sample of 401 toddlers from six districts in Central Papua. Data processing and analysis were conducted using RStudio software. Stunting status was determined based on the z-score of the height-for-age index (HAZ) calculated using the WHO Anthro software. The analysis results showed that out of 11 predictor variables, only age and height were significant predictors of stunting. The F1-score values for the Binary Logistic Regression method and the CART method are 95.73% and 90.38%, respectively. Thus, the Binary Logistic Regression method is more optimal in identifying factors influencing stunting in infants in Central Papua in 2023.

Keywords: Stunting, Binary Logistic Regression, CART, Classification, Central Papua.

BAB I

PENDAHULUAN

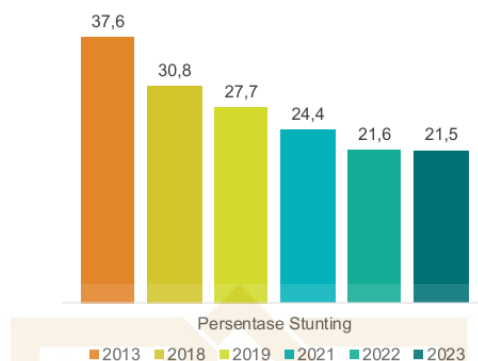
1.1 Latar Belakang

Kesehatan anak merupakan salah satu indikator penting dalam menilai kualitas kesehatan masyarakat secara keseluruhan, yang mencerminkan kualitas sumber daya manusia dan kesejahteraan generasi penerus. Indikator ini dapat berjalan optimal apabila anak mendapatkan asupan gizi yang cukup sejak dalam kandungan (Insyroh, 2020). Kekurangan gizi dapat menyebabkan berbagai masalah kesehatan, salah satunya adalah malnutrisi yang masih menjadi perhatian serius di berbagai negara, termasuk Indonesia.

Stunting merupakan salah satu bentuk malnutrisi kronis yang menghambat pertumbuhan linier anak. *Stunting* didefinisikan sebagai kondisi gagal tumbuh akibat kekurangan gizi kronis pada periode 1.000 Hari Pertama Kehidupan (HPK), ditandai dengan panjang atau tinggi badan yang kurang dari minus dua standar deviasi (-2 SD) dari median standar WHO (Kemenkes RI, 2023).

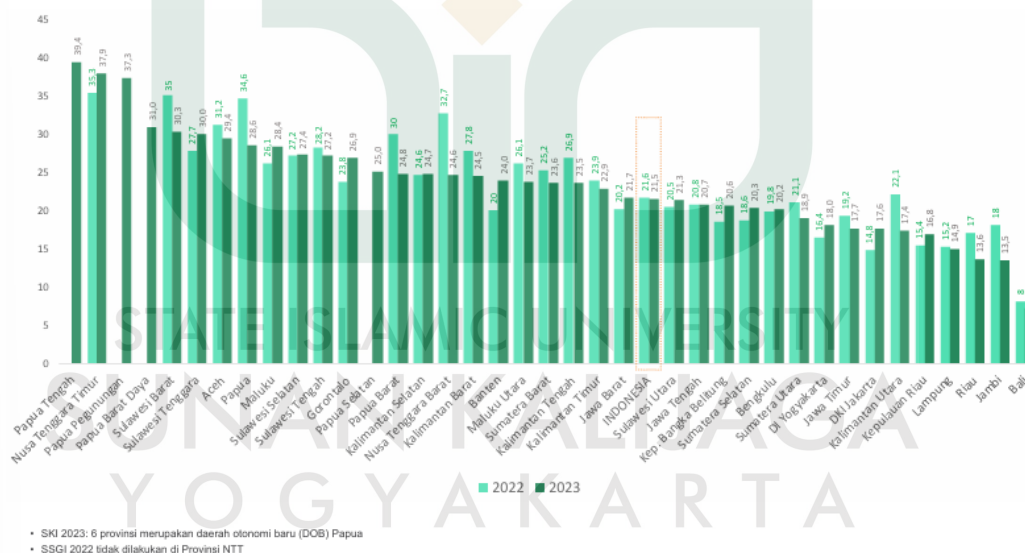
Dampak *stunting* tidak hanya terbatas pada kesehatan fisik, tetapi juga mengganggu perkembangan otak, yang berdampak pada kecerdasan. Anak *stunting* memiliki rata-rata skor *Intelligence Quotient* (IQ) 11 poin lebih rendah dibandingkan pada anak normal (Trihono et al., 2015). Dalam jangka panjang, *stunting* meningkatkan resiko berbagai penyakit tidak menular (PTM) di masa dewasa, seperti obesitas, hipertensi, diabetes hingga kanker. (Kemenkes RI, 2023).

Selain berdampak pada individu, *stunting* juga berdampak luas pada perekonomian negara. *Stunting* dapat menurunkan produktivitas angkatan kerja, memperbesar kesenjangan sosial ekonomi, serta menurunkan pertumbuhan ekonomi hingga 11% dari Produk Domestik Bruto (PDB) (Rahayu et al., 2018). Anak yang mengalami *stunting* berpeluang lebih kecil untuk menyelesaikan pendidikan tinggi dan mendapatkan pekerjaan dengan penghasilan layak, sehingga meningkatkan risiko kemiskinan antargenerasi.



Gambar 1.1 Persentase *stunting* di Indonesia tahun 2013-2023

Walaupun prevalensi *stunting* di Indonesia masih cukup tinggi, namun Indonesia sudah berhasil menurunkan prevalensi *stunting* dari 37,6% pada tahun 2013 menjadi 21,6% pada tahun 2022, dengan rata-rata penurunan sekitar 1,55% per tahun (Gambar 1.1). Prevalensi tersebut kemudian relatif stagnan pada temuan SKI 2023, yaitu di angka 21,5%. Perlu diakui progress ini belum dapat memenuhi target RPJMN 2020-2024 yang menargetkan prevalensi *stunting* sebesar 14% pada tahun 2024 (Kemenkes RI, 2023).



Gambar 1.2 Prevalensi *stunting* menurut provinsi Indonesia tahun 2022-2023

Hasil SKI 2023 menunjukkan bahwa sekitar 1 dari 5 balita usia 0-59 bulan di Indonesia mengalami *stunting* pada tahun 2023. Provinsi Papua Tengah mencatat prevalensi *stunting* tertinggi di Indonesia (Gambar 1.2), yaitu 39,4% dengan 46.128

kasus. Pemerintah Provinsi Papua Tengah menjadikan percepatan penurunan *stunting* sebagai salah satu dari tiga program prioritas di bidang kesehatan.

Tingginya prevalensi *stunting* dipengaruhi oleh berbagai faktor. *World Health Organization* (WHO) memiliki *conceptual framework* yang berisi empat faktor utama penyebab *stunting* pada balita, yaitu kondisi keluarga dan rumah tangga, kelengkapan makanan yang tidak memadai, praktik menyusui yang tidak optimal, serta infeksi. Kompleksitas penyebab ini memerlukan pendekatan analisis yang tepat untuk mengidentifikasi faktor-faktor dominan yang berpengaruh (Insyroh, 2020). Sejauh ini, penanganan *stunting* masih berfokus pada pengukuran langsung status gizi yang memerlukan biaya tinggi. Oleh karena itu, diperlukan pemanfaatan metode klasifikasi berbasis data *non-nutritional* seperti data usia, jenis kelamin, berat badan, dan tinggi badan untuk memperkirakan risiko *stunting* (Nababan, 2021).

Metode klasifikasi kini banyak diterapkan di berbagai bidang, termasuk pada bidang kesehatan. Klasifikasi dapat dilakukan dengan dua pendekatan metode yaitu dengan metode parametrik dan metode non-parametrik. Pada penelitian ini, metode klasifikasi yang digunakan adalah Regresi Logistik Biner sebagai metode parametrik dan *Classification and Regression Tree* (CART) sebagai metode non-parametriknya.

Regresi Logistik Biner adalah metode statistik yang memodelkan hubungan antara variabel respon biner (Y) dengan variabel prediktor (X) yang berskala kategorik ataupun rasio. Metode ini memiliki keunggulan dalam interpretasi yang sederhana dan kemampuan untuk menangani hubungan linier antar variabel. Selain itu, metode ini sering menghasilkan model dengan akurasi tinggi dalam klasifikasi data biner (Hosmer dan Lemeshow, 2000). Pada penelitian Aristawidya et al., (2024) menunjukkan bahwa metode regresi logistik biner memiliki ketepatan klasifikasi terbaik (92.31%) dalam memprediksi faktor resiko diabetes dibandingkan dengan *naïve bayes classifier* (84.61%).

Classification and Regression Tree (CART) merupakan metode non-parametrik yang menggambarkan hubungan antara variabel respon dengan satu atau lebih variabel prediktor yang bertujuan untuk mendapatkan kelompok data yang

akurat sebagai ciri dari suatu pengklasifikasian. Metode CART mampu menangani data dengan banyak variabel dan interaksi yang kompleks (Breiman et al., 1984). Penelitian Ikhsan et al., (2023) menunjukkan bahwa metode CART memberikan hasil klasifikasi terbaik (91,35%) dalam memprediksi penerima beasiswa dibandingkan dengan metode K-NN dan *Naïve Bayes*.

Berdasarkan latar belakang tersebut, penelitian ini memanfaatkan metode Regresi Logistik Biner dan *Classification and Regression Tree* (CART) dalam membandingkan ketepatan klasifikasi status *stunting* pada balita di Papua Tengah berdasarkan nilai *F1-score* serta mengidentifikasi variabel yang berpengaruh pada status *stunting*. Hasil penelitian ini diharapkan dapat memberikan informasi yang berguna bagi pemerintah dalam menentukan kebijakan penanganan masalah *stunting* di Indonesia, khususnya di Papua Tengah.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, dirumuskan beberapa rumusan masalah dalam penelitian ini yaitu:

1. Bagaimana algoritma metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART)?
2. Bagaimana penerapan metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART) pada data status *stunting* balita di Papua Tengah tahun 2023?
3. Bagaimana menentukan variabel yang signifikan terhadap status *stunting* pada balita di Papua Tengah tahun 2023 dengan menggunakan metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART)?
4. Bagaimana perbandingan ketepatan klasifikasi metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART) berdasarkan nilai *F1-score*?

1.3 Batasan Masalah

Adapun batasan masalah yang dilakukan pada penelitian ini agar lebih terarah, yaitu sebagai berikut:

1. Penerapan metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART) hanya dilakukan pada data balita di Papua Tengah dengan menggunakan dataset Survei Kesehatan Indonesia (SKI) 2023 yang diperoleh dari Pusat Data dan Teknologi Informasi (PUSDATIN) Kemenkes RI. Variabel prediktor yang digunakan dibatasi pada variabel-variabel yang tersedia dalam dataset dan telah melewati proses seleksi awal.
2. Perbandingan ketepatan klasifikasi antara metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART) dilakukan berdasarkan nilai *F1-Score* sebagai ukuran evaluasi utama.
3. Alat pengolahan data menggunakan *software* RStudio 4.4.2 dan WHO Anthro 3.2.2.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, tujuan dalam penelitian ini adalah sebagai berikut:

1. Menjelaskan algoritma metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART).
2. Menerapkan metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART) pada data status *stunting* balita di Papua Tengah tahun 2023.
3. Menentukan variabel-variabel yang signifikan terhadap status *stunting* balita di Papua Tengah tahun 2023 berdasarkan hasil analisis menggunakan metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART).
4. Membandingkan hasil ketepatan klasifikasi antara metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART) dalam klasifikasi status *stunting* pada balita di Papua Tengah tahun 2023 berdasarkan nilai *F1-score*.

1.5 Manfaat Penelitian

Berdasarkan hasil dari penelitian ini, diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Teoritis

Menambah referensi dalam penerapan Regresi Logistik Biner dan *Classification and Regression Tree* (CART) dalam klasifikasi status *stunting* serta menjadi bahan kajian bagi peneliti lain yang tertarik dalam analisis klasifikasi pada penelitian serupa.

2. Manfaat Praktis

Memberikan informasi bagi tenaga kesehatan dan pemangku kebijakan mengenai faktor yang berpengaruh terhadap *stunting* berdasarkan analisis metode yang digunakan, serta membantu dalam pemilihan metode analisis yang lebih akurat untuk klasifikasi status *stunting*.

1.6 Tinjauan Pustaka

Tinjauan pustaka dalam penelitian ini berisi tentang literatur yang relevan dengan topik pembahasan dari penelitian yang sudah ada untuk dijadikan rujukan penelitian. Berikut beberapa rujukan utama di dalam bab ini:

1. Penelitian yang dilakukan oleh Rafika Aristawidya, Indahwati, Erfiani, Anwar Fitrianto, dan Muftih A. A. (2024) dengan judul “Perbandingan Analisis Regresi Logistik Biner dan *Naïve Bayes Classifier* untuk Memprediksi Faktor Resiko Diabetes”. Penelitian ini menjelaskan mengenai kinerja metode regresi logistik biner dan *naïve bayes classifier* dalam memprediksi risiko diabetes serta mengidentifikasi faktor-faktor yang secara signifikan mempengaruhi risiko diabetes. Hasil dari penelitian ini menunjukkan bahwa faktor yang berpengaruh terhadap risiko diabetes yaitu *age* (usia), *gender* (jenis kelamin), *polyuria* (frekuensi buang air kecil meningkat), *polydipsia* (rasa haus berlebih), *genital thrush* (infeksi jamur pada alat kelamin), *itching* (gatal-gatal), *irritability* (cepat marah), dan *partial paresis* (kelemahan sebagian otot). Selanjutnya kedua metode dievaluasi berdasarkan akurasi, sensitivitas, spesifisitas, dan nilai prediktif positif. Hasil regresi logistik biner memiliki ketepatan klasifikasi lebih

tinggi (92.31%) dibandingkan dengan *naïve bayes classifier* (84.61%). Oleh karena itu, regresi logistik biner diidentifikasi sebagai metode terbaik untuk memprediksi risiko diabetes dalam konteks penelitian ini.

2. Penelitian yang dilakukan oleh Yasmine Nurfirdaus, Rery Afianto, dan Aldiella Ayu Prasetya Astuti (2024) dengan judul “Klasifikasi Faktor Ibu terhadap Risiko Terjadinya *Stunting* Menggunakan Metode *Classification and Regression Tree* (CART) (Studi Pada Balita Usia 0-24 Bulan di Kec. Bluluk, Kab. Lamongan)”. Penelitian ini menjelaskan mengenai analisis pengaruh faktor ibu, termasuk pemberian ASI eksklusif, makanan pendamping ASI (MP-ASI), dan pola asuh, terhadap risiko *stunting* menggunakan metode *Classification and Regression Tree* (CART). Dalam penelitian ini, metode CART dievaluasi berdasarkan nilai akurasi dan nilai APER (*Apparent Error Rate*). Hasil dari penelitian ini menunjukkan bahwa faktor yang paling berpengaruh terhadap risiko *stunting* adalah pemberian ASI eksklusif dengan nilai ketepatan klasifikasi sebesar 79,7%.
3. Penelitian yang dilakukan oleh Lulu Anata Sanur, M. Al Haris, dan Fatkhurokhman Fauzi (2024) dengan judul “Klasifikasi Status Kesejahteraan Masyarakat Kabupaten Kepulauan Mentawai dengan Metode Regresi Logistik Biner dan *Classification and Regression Tree* (CART)”. Penelitian ini menjelaskan mengenai pengaruh karakteristik atau variabel pada status kesejahteraan rumah tangga menggunakan metode regresi logistik biner dan CART. Dalam penelitian ini, kedua metode dievaluasi berdasarkan nilai akurasi dan nilai APER (*Apparent Error Rate*). Hasil dari penelitian ini menunjukkan bahwa variabel usia, tingkat pendidikan terakhir, dan kesehatan kepala rumah tangga memiliki pengaruh signifikan terhadap model klasifikasi. Selanjutnya, hasil regresi logistik biner memiliki ketepatan klasifikasi lebih tinggi (93,94%) dibandingkan dengan CART (89,40%). Oleh karena itu, model terbaik yang digunakan dalam penelitian ini adalah metode regresi logistik biner.
4. Penelitian yang dilakukan oleh Ali Nur Ikhsan, Pungkas Subarkah, dan Raditya Sani Alifian (2023) dengan judul “Komparasi Algoritme K-NN, *Naïve Bayes*, dan CART untuk Memprediksi Penerima Beasiswa”. Penelitian ini menjelaskan

mengenai analisis komparasi algoritme K-NN, Naïve Bayes, dan CART untuk memprediksi penerima beasiswa bagi pengelola di SMA. Hasil perbandingan ketiga algoritme mendapatkan nilai hasil yaitu algoritme K-NN sebesar 90.0844%, algoritme *Naïve Bayes* sebesar 90.8228%, dan algoritme CART sebesar 91.3502%. Oleh karena itu, dapat diketahui bahwa dari ketiga algoritme menunjukkan bahwa algoritme CART memperoleh nilai akurasi tertinggi yang didasarkan pada nilai *Precision*, *Recall* dan *F-Measure* dengan kategori *Good Classification*.

5. Penelitian yang dilakukan oleh Elfira Safitri, Sri Basriati, dan Septia Mulyani (2022) dengan judul “Faktor-Faktor yang Mempengaruhi Kejadian *Stunting* terhadap Balita Menggunakan Analisis Regresi Logistik”. Penelitian ini menjelaskan tentang faktor-faktor yang mempengaruhi kejadian *stunting* pada balita di Posyandu Kasih Ibu Puskesmas Airtiris, Kampar menggunakan metode regresi logistik biner. Hasil dari penelitian ini menunjukkan bahwa faktor-faktor yang mempengaruhi *stunting*, yaitu status gizi berat badan berdasarkan usia. Selanjutnya, nilai ketepatan klasifikasi model dihitung berdasarkan nilai akurasi yaitu sebesar 83,6%.

Berikut persamaan dan perbedaan penelitian ini dengan penelitian yang telah dipaparkan sebelumnya.

Tabel 1.1 Tinjauan pustaka

Judul	Hasil	Persamaan	Perbedaan
Perbandingan Analisis Regresi Logistik Biner dan <i>Naïve Bayes Classifier</i> untuk Memprediksi Faktor Resiko Diabetes (Aristawidya et al., 2024)	Regresi logistik biner memiliki nilai ketepatan klasifikasi lebih tinggi (92,31%) dibandingkan dengan <i>naïve bayes classifier</i> (84.61%).	Penggunaan metode regresi logistik dalam penelitian khususnya di bidang kesehatan.	Tidak menggunakan metode CART sebagai perbandingan hasil ketepatan klasifikasi.

Klasifikasi Faktor Ibu terhadap Risiko Terjadinya <i>Stunting</i> menggunakan Metode <i>Classification and Regression Tree</i> (CART) (Studi Pada Balita Usia 0-24 Bulan di Kec. Bluluk, Kab. Lamongan) (Nurfirdaus et al., 2024)	Metode <i>Classification and Regression Tree</i> (CART) menghasilkan ketepatan klasifikasi sebesar 79,7% dengan faktor yang paling berpengaruh yaitu pemberian ASI eksklusif.	Penggunaan metode <i>Classification and Regression Tree</i> (CART) dalam kasus <i>stunting</i> .	Tidak membandingkan dengan metode regresi logistik biner.
Klasifikasi Status Kesejahteraan Masyarakat Kabupaten Kepulauan Mentawai dengan Metode Regresi Logistik Biner dan <i>Classification and Regression Tree</i> (CART) (Sanur et al., 2024)	Regresi logistik biner memiliki nilai ketepatan klasifikasi lebih tinggi (93,94%) dibandingkan dengan CART (89,40%).	Penggunaan metode regresi logistik biner dan <i>Classification and Regression Tree</i> (CART).	Hasil perbandingan yang dilihat dari perhitungan ketepatan klasifikasi berdasarkan nilai akurasi dan nilai APER (<i>Apparent Error Rate</i>).
Komparasi Algoritme K-NN, Naïve Bayes, dan CART untuk	Algoritme CART memiliki hasil akurasi yang paling tinggi yaitu sebesar	Penggunaan metode CART dalam konteks klasifikasi.	Tidak menggunakan metode regresi logistik biner

Memprediksi Penerima Beasiswa (Ikhsan et al., 2023)	91,35% dengan kategori <i>Good Classification</i> .		sebagai perbandingan hasil ketepatan klasifikasi.
Faktor-Faktor yang Mempengaruhi Kejadian <i>Stunting</i> terhadap Balita Menggunakan Analisis Regresi Logistik (Safitri et al., 2022)	Regresi logistik biner memiliki nilai ketepatan klasifikasi sebesar 83,6% dengan faktor yang mempengaruhi adalah status gizi berat badan berdasarkan usia.	Penggunaan metode regresi logistik biner dalam kasus <i>stunting</i> .	Tidak menggunakan metode CART sebagai perbandingan hasil ketepatan klasifikasi.

1.7 Sistematika Penulisan

Sistematika penulisan bertujuan untuk memudahkan membaca dalam memahami tulisan dalam penelitian ini secara sederhana, rinci, runtut dan jelas. Adapun sistematika penulisan dalam penelitian ini adalah sebagai berikut:

1. BAB I PENDAHULUAN

Pada bab ini berisi tentang hal yang melatarbelakangi peneliti dalam melakukan suatu penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan tinjauan pustaka yang mana menjadi acuan penelitian dalam penyusunan, serta sistematika penulisan sebagai gambaran sederhana dari penelitian ini secara garis besar.

2. BAB II LANDASAN TEORI

Pada bab ini berisi mengenai teori dasar dan teori penunjang sebagai penguat pembahasan analisis dalam klasifikasi status *stunting* pada balita dengan metode Regresi Logistik Biner dan *Classification and Regression Tree* (CART).

3. BAB III METODE PENELITIAN

Pada bab ini berisi mengenai jenis penelitian, jenis dan sumber data, variabel penelitian, alat pengolahan data, metode analisis data, dan *flowchart* dari penelitian ini.

4. BAB IV PEMBAHASAN

Pada bab ini berisi tentang metode analisis yang diteliti, sampai dengan pembahasan terperinci mengenai metode Regresi Logistik Biner dan *Classification and Regression Tree* (CART).

5. BAB V STUDI KASUS

Pada bab ini berisi tentang penerapan metode Regresi Logistik Biner dan *Classification and Regression Tree* (CART) dalam klasifikasi status *stunting* pada balita di Papua Tengah tahun 2023 menggunakan *Software* RStudio 4.4.2 dan WHO Anthro 3.2.2.

6. BAB VI PENUTUP

Pada bab ini berisi tentang kesimpulan yang dihasilkan dari pembahasan dan jawaban atas rumusan masalah yang telah dirancang sebelumnya, serta berisikan saran bagi peneliti selanjutnya.

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil dari penelitian tentang klasifikasi status *stunting* pada balita di Papua Tengah tahun 2023, dapat diperoleh kesimpulan sebagai berikut:

1. Algoritma metode Regresi Logistik Biner dan *Classification and Regression Tree* (CART) yaitu:
 - Regresi Logistik Biner menggunakan pendekatan fungsi logit untuk memodelkan hubungan antara variabel prediktor dengan probabilitas terjadinya *stunting*. Proses utama dalam metode ini meliputi pengujian multikolinearitas, estimasi parameter, pengujian signifikansi parameter, pembentukan model, serta interpretasi model.
 - *Classification and Regression Tree* (CART) bekerja dengan membentuk pohon keputusan melalui proses pemilihan pemilah, penentuan simpul terminal, dan penandaan label kelas. Selanjutnya, melakukan pemangkasan pohon untuk mendapatkan pohon klasifikasi yang optimal.
2. Penerapan metode Regresi Logistik Biner dan metode *Classification and Regression Tree* (CART) yaitu:
 - Metode Regresi Logistik Biner membentuk model yang diperoleh dalam bentuk logit sebagai berikut:
$$g(x) = 40.03190 + 0.44187X_3 - 0.64265X_5$$
 - Metode *Classification and Regression Tree* (CART) membentuk pohon klasifikasi optimal yang disajikan pada Gambar 5.10 menghasilkan 5 simpul induk dan 6 simpul terminal dengan tingkat kedalaman sebesar 3 tingkatan. Variabel yang paling menentukan klasifikasi status *stunting* adalah variabel tinggi badan (X_5) dan usia (X_3).
3. Hasil analisis dari metode Regresi Logistik Biner dan metode CART mengatakan bahwa dari 11 variabel prediktor, hanya variabel usia (X_3) dan tinggi badan (X_5) yang signifikan terhadap status *stunting*.

4. Perbandingan ketepatan klasifikasi dari kedua metode berdasarkan nilai *F1-score* menunjukkan bahwa metode Regresi Logistik Biner memiliki nilai ketepatan klasifikasi sebesar 95.73% sedangkan metode *Classification and Regression Tree* (CART) sebesar 90.38%. Hal ini dapat disimpulkan bahwa metode regresi logistik biner lebih unggul dalam hal ketepatan klasifikasi dengan 5.35% lebih tinggi dibandingkan dengan metode CART.

6.2 Saran

Berdasarkan dari hasil penelitian, berikut merupakan beberapa saran yang dapat digunakan untuk penelitian selanjutnya yaitu:

1. Mengeksplorasi metode atau pendekatan lain yang dapat menghasilkan model dengan kelayakan yang lebih baik, namun tetap mempertahankan atau bahkan meningkatkan kemampuan klasifikasinya.
2. Pemilihan variabel prediktor yang lebih relevan dan lebih lengkap terkait faktor penyebab balita *stunting*, seperti pola makan, ketahanan pangan keluarga, atau status keluarga.
3. Berdasarkan variabel yang signifikan dalam model, pemerintah dapat merancang intervensi gizi dan edukasi yang lebih tepat sasaran, terutama untuk wilayah dengan prevalensi tinggi.

DAFTAR PUSTAKA

- Agresti, A. (2013). *Categorical Data Analysis* (Third edition). John Wiley & Sons.
- Anton, H., & Rorres, C. (2013). *Elementary linear algebra: applications version*. John Wiley & Sons.
- Aprilia, D., & Tono, S. F. N. (2023). Pengaruh Status Imunisasi Dasar Terhadap Kejadian Stunting dan Gangguan Perkembangan Balita. *Jurnal Kebidanan*, 12(1), 66–74.
- Aribowo, B., & Fairuz, S. (2024). *Panduan Praktis Machine Learning Klasifikasi Menggunakan Python: Diandra Kreatif*. Diandra Kreatif.
- Aristawidya, R., Indahwati, Erfiani, Fitrianto, A., & Muftih, A. A. (2024). Perbandingan Analisis Regresi Logistik Biner Dan Naïve Bayes Classifier untuk Memprediksi Faktor Resiko Diabetes. *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 5(2), 782–794. <https://doi.org/https://doi.org/10.46306/lb.v5i2.617>
- Astari, L. D., Nasoetion, A., & Dwiriani, C. M. (2005). *Hubungan karakteristik keluarga, pola pengasuhan dan kejadian stunting anak usia 6-12 bulan*.
- Bardosono, S., Sastroamidjojo, S., & Lukito, W. (2007). Determinants of child malnutrition during the 1999 economic crisis in selected poor areas of Indonesia. *Asia Pacific Journal of Clinical Nutrition*, 16(3).
- Batiro, B., Demissie, T., Halala, Y., & Anjulo, A. A. (2017). Determinants of stunting among children aged 6-59 months at Kindo Didaye woreda, Wolaita Zone, Southern Ethiopia: Unmatched case control study. *PLoS ONE*, 12(12).
- Breiman, L., Friedman, J. H., Olshen, R., & Stone, C. J. (1984). *Classification and Regression Trees*.
- Deepchecks. (2024). *Understanding F1 Score, Accuracy, ROC-AUC, and PR-AUC Metrics for Models*. Deepchecks Community Blog. <https://www.deepchecks.com/f1-score-accuracy-roc-auc-and-pr-auc-metrics-for-models/>. Diakses tanggal 01 Maret 2025.
- Encord. (2023, July 18). *F1 score in machine learning*. Encord Blog. <https://encord.com/blog/f1-score-in-machine-learning/>. Diakses tanggal 01 Maret 2025.
- Fantay Gebru, K., Mekonnen Haileselassie, W., Haftom Temesgen, A., Oumer Seid, A., & Afework Mulugeta, B. (2019). Determinants of stunting among under-five children in Ethiopia: a multilevel mixed-effects analysis of 2016 Ethiopian demographic and health survey data. *BMC Pediatrics*, 19, 1–13.
- Gorunescu, F. (2011). *Data Mining: Concepts, Models and Techniques* (Vol. 12). Springer Science & Business Media. <https://doi.org/10.1007/978-3-642-19721-5>

- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (Second Edition). John Wiley & Sons, Inc.
- Ikhsan, A. N., Subarkah, P., & Alifian, R. S. (2023). Komparasi Algoritme K-NN, Naïve Bayes, dan CART untuk Memprediksi Penerima Beasiswa. *JST (Jurnal Sains Dan Teknologi)*, 12(2), 309–316. <https://doi.org/https://doi.org/10.23887/jstundiksha.v12i2.51745>
- Insyroh, W. N. (2020). Klasifikasi Balita Terkait Kasus Stunting di Kota Surabaya Menggunakan Classification and Regression Trees dengan Adaptive Synthetic Nominal dan Adaptive Synthetic K-Nearest Neighbor. *Skripsi*. Surabaya: Institut Teknologi Sepuluh Nopember.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, No. 1). New York: springer.
- Johnson, R. A., Miller, I., & Freund, J. E. (2000). *Probability and statistics for engineers*. Prentice Hall.
- Kemendes RI. (2020). *Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 tentang Standar Antropometri Anak*. Jakarta: Kementerian Kesehatan RI.
- Kemendes RI. (2023). *Survei Kesehatan Indonesia Tahun 2023 Dalam Angka*. In Jakarta: BPPK, Kemendes RI.
- Kemendes RI. (2023). *BAB 4: Stunting di Indonesia dan Faktor Determinan*.
- Kollo, T. (2005). *Advanced multivariate statistics with matrices*. Springer.
- Lewis, R. J. (2000). An introduction to classification and regression tree (CART) analysis. *Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California*, 14.
- Maulid, R. (2023). *Pahami aturan 80/20 dalam data science*. DQLab. <https://dqlab.id/pahami-aturan-80/20-dalam-data-science>. Diakses tanggal 05 Januari 2025.
- Muflikhah, L., Ratnawati, D. E., & MP, R. R. (2018). *Data mining*. Universitas Brawijaya Press.
- Nababan, J. F. (2021). *Klasifikasi Penderita Stunting Dengan Metode Support Vector Machine (Studi Kasus: Lima Puskesmas Di Kota Bandar Lampung)*.
- Noeryanti, N. (2021). *Pengantar Teori Probabilitas*. AKPRIND PRESS.
- Nurfirdaus, Y., Afianto, R., & Astuti, A. A. P. (2024). Klasifikasi Faktor Ibu terhadap Risiko Terjadinya Stunting menggunakan Metode Classification and Regression Tree (CART) (Studi Pada Balita Usia 0-24 Bulan di Kec. Bluluk, Kab. Lamongan). *PREPOTIF: JURNAL KESEHATAN MASYARAKAT*, 8(3), 6209–6216.
- Qudratullah, M. F. (2013). *Analisis Regresi Terapan: Teori, Contoh Kasus, dan Aplikasi dengan SPSS*. Penerbit Andi.

- Rahayu, A., Yulidasari, F., Putri, A. O., & Anggraini, L. (2018). Study guide-stunting dan upaya pencegahannya. *Buku Stunting Dan Upaya Pencegahannya*, 88.
- Roflin, E., Riana, F., Munarsih, E., Pariyana, & Liberty, I. A. (2023). *Regresi Logistik Biner Dan Multinomial* (Nasrudin, Ed.). Penerbit NEM.
- 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.
- Sanur, L. A., Al Haris, M., & Fauzi, F. (2024). Klasifikasi Status Kesejahteraan Masyarakat Kabupaten Kepulauan Mentawai dengan Metode Regresi Logistik Biner dan Classification and Regression Tree (CART). *PARAMETER: Jurnal Matematika, Statistika Dan Terapannya*, 3(01), 71–84. <https://doi.org/https://doi.org/10.30598/parameterv3i01pp71-84>.
- Sumartini, S. H., & Purnami, S. W. (2016). Penggunaan metode classification and regression trees (CART) untuk klasifikasi rekurensi pasien kanker serviks di RSUD Dr. Soetomo Surabaya. *Jurnal Sains Dan Seni ITS*, 4(2).
- Suratri, M. A. L., Putro, G., Rachmat, B., Nurhayati, Ristrini, Pracoyo, N. E., Yulianto, A., Suryatma, A., Samsudin, M., & Raharni. (2023). Risk factors for *stunting* among children under five years in the province of East Nusa Tenggara (NTT), Indonesia. *International Journal of Environmental Research and Public Health*, 20(2), 1640.
- Susanto, S., & Suryadi, D. (2010). *Pengantar data mining: mengagali pengetahuan dari bongkahan data*. Penerbit Andi.
- Tiara Aprilia, K., Gusriani, N., & Parmikanti, K. (2015). Klasifikasi Ketepatan Masa Studi Mahasiswa FMIPA Unpad Angkatan 2001-2006 dengan Menggunakan Metode Classification and Regression Trees (CART). *Jurnal Matematika Integratif*, 11, 7–14.
- Titaley, C. R., Ariawan, I., Hapsari, D., Muasyaroh, A., & Dibley, M. J. (2019). Determinants of the stunting of children under two years old in Indonesia: A multilevel analysis of the 2013 Indonesia basic health survey. *Nutrients*, 11(5), 1106.
- Trihono, T., Atmarita, A., Tjandrarini, D. H., Irawati, A., Nurlinawati, I., Utami, N. H., & Tejayanti, T. (2015). *Pendek (Stunting) di Indonesia, Masalah dan Solusinya*. Lembaga Penerbit Badan Litbangkes.
- Widana, I. W. & Muliani, N. P. L. (2020). *Uji persyaratan analisis*.
- World Health Organization. (2023). *Leadership Dialogue on Food Systems for People's Nutrition and Health*. World Health Organization. <https://www.who.int/news/item/28-07-2023-leadership-dialogue-on-food-systems-for-people-s-nutrition-and-health>. Diakses pada 10 Oktober 2024.