

**KOMPARASI ALGORITMA DECISION TREE C4.5 DAN  
NAIVE BAYES PADA KLASIFIKASI STATUS GIZI BALITA  
STUNTING**

**SKRIPSI**



**Oleh:**

**ISHAK FEBRIANTO**

**NPM.19081010072**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR  
2023**

# **KOMPARASI ALGORITMA DECISION TREE C4.5 DAN NAIVE BAYES PADA KLASIFIKASI STATUS GIZI BALITA STUNTING**

## **SKRIPSI**

Diajukan Untuk Memenuhi Sebagai Persyaratan Dalam Menempuh Gelar Sarjana  
Komputer Program Studi Informatika



Oleh:

**ISHAK FEBRIANTO**

**NPM.19081010072**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR  
2023**

## LEMBAR PENGESAHAN

### SKRIPSI

Judul : KOMPARASI ALGORITMA DECISION TREE C4.5 DAN  
NAIVE BAYES PADA KLASIFIKASI STATUS GIZI BALITA  
STUNTING

Oleh : Ishak Febrianto

NPM : 19081010072

Telah Diseminarkan Dalam Ujian Skripsi Pada:

Hari Senin, Tanggal 22 Mei 2023

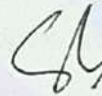
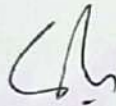
Mengetahui:

Dosen Pembimbing

Dosen Penguji

1.

1.

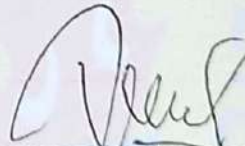


Eva Yulia Puspaningrum,  
S.Kom, M.Kom  
NIP. 19890705 2021212 002

Eva Yulia Puspaningrum,  
S.Kom, M.Kom  
NIP. 19890705 2021212 002

2.

2.



Dr. Eng. Ir. Anggraini Puspita  
Sari, S.T., M.T.  
NPT. 222198 60 816400

Muhammad Muharrôm Al  
Haromainy, S.Kom, M.Kom  
NIP. 19950601 202203 1000

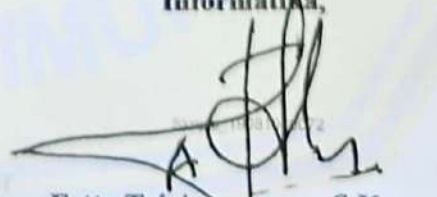
Menyetujui:

Dekan  
Fakultas Ilmu Komputer

Koordinator Program Studi  
Informatika,



Dr. Novirina Hendrasarie,  
M.T.  
NIP. 19681126 199403 2 001



Fetty Tri Anggraeny, S.Kom.,  
M.Kom.  
NIP. 19820211 2021212 005

## **SURAT PERNYATAAN ANTI PLAGIAT**

Saya, mahasiswa program studi Informatik UPN “Veteran” Jawa Timur, yang bertanda tangan di bawah ini:

Nama : ISHAK FEBRIANTO

NPM : 19081010072

Menyatakan bahwa judul skripsi yang saya ajukan dan kerjakan yang berjudul

**“KOMPARASI ALGORITMA DECISION TREE C4.5 DAN NAIVE BAYES PADA KLASIFIKASI STATUS GIZI BALITA STUNTING ”**

Bukan merupakan plagiat dari skripsi/tugas akhir/penelitian orang lain dan juga bukan merupakan produk dan atau software yang saya beli dari pihak lain. Saya juga menyatakan bahwa skripsi ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam daftar pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di UPN “Veteran” Jawa Timur maupun di institusi Pendidikan lain.

Jika ternyata di kemudian hari pernyataan ini terbukti tidak benar, maka saya siap menerima segala konsekuensinya.

Surabaya, 10 Mei 2023

Penulis



**ISHAK FEBRIANTO**

**NPM. 19081010072**

# KOMPARASI ALGORITMA DECISION TREE C4.5 DAN NAIVE BAYES PADA KLASIFIKASI STATUS GIZI BALITA STUNTING

Nama Mahasiswa : Ishak Febrianto  
NPM : 19081010072  
Program Studi : Informatika  
Dosen Pembimbing : Eva Yulia Puspaningrum, S.Kom, M.Kom.  
Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT.

---

## ABSTRAK

Masalah gizi masih menjadi masalah yang krusial saat ini. Salah satu dari banyaknya masalah gizi yang ada yaitu *stunting*. *Stunting* menurut WHO (2015) adalah keadaan malnutrisi atau kurang gizi pada masa pertumbuhan dan perkembangan anak, yang ditandai dengan panjang atau tinggi badannya di bawah standar. *Data mining* dengan teknik klasifikasi pada status gizi balita *stunting* dapat dilakukan untuk membantu dalam mengidentifikasi balita yang mengalami *stunting* dan memberikan pengukuran yang objektif tentang status gizi mereka.

Pada penelitian ini akan dibandingkan metode algoritma *decision tree* C4.5 yang termasuk dalam pendekatan pohon keputusan dan *naïve bayes* yang menggunakan pendekatan berbasis probabilitas kemunculan kelas dalam mengklasifikasikan status gizi balita *stunting* dengan dilakukan *discretization* pada tahap *preprocessing*. Data yang digunakan dalam penelitian ini didapatkan dari Puskesmas Jagir, Surabaya berupa data sekunder gizi balita tahun 2021 sebanyak 2801. Pelabelan *stunting* atau normal pada dataset menggunakan acuan standar antropometri anak di Indonesia yang tercantum pada Peraturan Menteri Kesehatan Republik Indonesia nomor 2 tahun 2020. Apabila nilai z-score antropometri anak kurang dari minus 2 standar deviasi (SD) berdasarkan indeks panjang badan menurut umur (PB/U) atau tinggi badan menurut umur (TB/U) maka anak termasuk dalam indikasi *stunting*.

Perbandingan model dianalisis dari 3 skenario pembagian data latih dan uji yang dilakukan (70:30,80:20,90:10). Evaluasi model klasifikasi dianalisis menggunakan metrik akurasi, presisi, *recall*, *f1-score*, dan *AUC score* pada grafik ROC pada setiap skenarionya. Didapatkan hasil pada skenario terbaik dengan perbandingan data latih dan uji 70 : 30 bahwa *decision tree* C4.5 lebih unggul berdasarkan *AUC score* sebesar 86% (*good classification*) sedangkan *naïve bayes* menghasilkan *AUC score* sebesar 77% (*fair classification*) dalam mengklasifikasikan status gizi balita *stunting*.

**Kata kunci:** Klasifikasi, *Decision Tree* C4.5, *Naïve bayes*, *Stunting*, *Data mining*.

## KATA PENGANTAR

Puji syukur kehadiran Tuhan YME atas rahmat-Nya yang telah melimpahkan segala nikmat kepada penulis sehingga dapat menyelesaikan skripsi ini dengan judul "Komparasi Algoritma Decision Tree C4.5 dan Naive Bayes pada Klasifikasi Status Gizi Balita Stunting". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Informatika.

Penulis ingin menyampaikan penghargaan dan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu dan mendukung penulis dalam menyelesaikan skripsi ini. Terima kasih kepada dosen pembimbing yang telah memberikan bimbingan, arahan, dan motivasi selama penulis menulis skripsi ini. Terima kasih juga kepada rekan-rekan yang telah memberikan dukungan, saran, dan masukan yang berarti dalam penulisan skripsi ini.

Skripsi ini membahas tentang klasifikasi status gizi balita stunting menggunakan algoritma Decision Tree C4.5 dan Naive Bayes. Penelitian ini dilakukan untuk membandingkan kinerja kedua algoritma dalam melakukan klasifikasi status gizi balita stunting. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari puskesmas Jagir, Surabaya.

Penulis berharap bahwa skripsi ini dapat memberikan manfaat dan kontribusi yang berguna bagi pembaca, terutama bagi para peneliti dan praktisi di bidang kesehatan serta teknologi informasi. Penulis juga menyadari bahwa skripsi ini masih memiliki kekurangan dan kelemahan, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk perbaikan penelitian di masa yang akan datang. Akhir kata, penulis berharap bahwa skripsi ini dapat diterima dan bermanfaat.

## UCAPAN TERIMA KASIH

Laporan skripsi ini dapat terselesaikan dengan baik atas anugerah dan rahmat Tuhan YME. Adapun dalam proses penyusunan skripsi, penulis mendapatkan bantuan, arahan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang tulus kepada Tuhan YME yang telah memberikan kelancaran, kemudahan, kekuatan, dan inspirasi kepada penulis. Penulis juga ingin mengucapkan terima kasih yang setinggi-tingginya kepada kedua orang tua dan nenek yang selalu memberikan dukungan penuh dalam setiap tahap penulisan skripsi ini. Sebelum melanjutkan, penulis ingin menyampaikan rasa terima kasih kepada:

1. Prof. Dr. Ir. Akhmad Fauzi, MMT., selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Dr. Ir. Novirina Hendrasarie, MT. Selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Fetty Tri Anggraeny, S.Kom, M.Kom selaku Koordinator Program Studi Teknik Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Eva Yulia Puspaningrum, S.Kom, M.Kom. selaku dosen pembimbing 1 skripsi ini yang selalu membimbing dengan baik, memberikan masukan dan saran-saran penelitian dengan sepenuh hati sehingga keseluruhan rangkaian penyusunan skripsi ini selesai tepat pada waktunya.
5. Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT. selaku dosen pembimbing 2 skripsi ini yang selalu memotivasi dan memberikan banyak arahan berharga dalam penulisan sehingga penyusunan skripsi ini dapat selesai dengan tepat waktu.
6. Seluruh anggota dosen pengajar yang tidak dapat saya sebutkan satu per satu, yang telah membimbing dan memberikan ilmu yang bermanfaat kepada penulis selama masa studi.
7. Seluruh staff Fakultas Ilmu Komputer UPN Veteran Jawa Timur yang telah membantu dalam penyelesaian rangkaian skripsi ini.
8. Poli Gizi Puskesmas Jagir Surabaya yang telah memperbolehkan, membimbing, dan membagi ilmu kepada penulis dalam melakukan penelitian.

9. Teman-teman Informatika angkatan 2019, yang selalu membantu, mengisi hari-hari, dan menjadi bagian dalam cerita hidup perkuliahan penulis.

10. Semua pihak yang tidak dapat disebutkan oleh penulis satu per satu.

Penulis mengucapkan terima kasih yang tulus atas semua bantuan yang telah diberikan. Semoga Allah SWT memberikan balasan yang sepadan kepada semua pihak yang telah membantu penulis dalam menyelesaikan penelitian dan menyusun laporan skripsi dengan baik dan tepat waktu.

Surabaya, 11 Mei 2023

Ishak Febrianto



## DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI.....	i
SURAT PERNYATAAN ANTI PLAGIAT .....	ii
ABSTRAK .....	iii
KATA PENGANTAR .....	iv
UCAPAN TERIMA KASIH.....	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR .....	ix
DAFTAR TABEL.....	x
BAB I PENDAHULUAN.....	1
1.1    Latar Belakang.....	1
1.2    Rumusan Masalah.....	5
1.3    Tujuan Penelitian .....	5
1.4    Manfaat Penelitian .....	5
1.5    Batasan Masalah .....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1.    Penelitian Terdahulu .....	7
2.2.    Balita.....	9
2.3.    Stunting.....	9
2.4.    Indeks Antropometri .....	10
2.5.    Puskesmas .....	12
2.6.    Data Mining .....	13
2.7. <i>Decision Tree</i> .....	14
2.8.    Algoritma C4.5 .....	15
2.9. <i>Naive Bayes Classifier</i> .....	16
2.10. <i>Confusion Matrix</i> .....	18

BAB III METODOLOGI PENELITIAN.....	21
3.1.    Prosedur penelitian.....	21
3.1.1.    Pengumpulan Data .....	23
3.1.2. <i>EDA (Exploratory Data Analysis)</i> .....	24
3.1.3. <i>Preprocessing Data</i> .....	24
3.1.4.    Pembagian Data.....	29
3.1.5.    Pemodelan .....	30
3.1.6.    Evaluasi Model.....	36
3.2.    Skenario Uji Coba.....	38
BAB IV HASIL DAN PEMBAHASAN .....	39
4.1    Proses Pengambilan Data.....	39
4.2 <i>EDA (Exploratory Data Analysis)</i> .....	39
4.3 <i>Preprocessing</i> .....	41
4.3.1    Penanganan <i>missing data</i> .....	41
4.3.2 <i>Label Encoder</i> .....	42
4.3.3.    SMOTE .....	43
4.3.4. <i>Discretization / Binning</i> .....	44
4.3.5. <i>Undersampling</i> .....	45
4.4    Skenario Uji Coba.....	47
4.4.1    Skenario 1 .....	47
4.4.2    Skenario 2 .....	63
4.4.3    Skenario 3.....	65
4.5    Hasil dan Analisis Uji.....	66
BAB V PENUTUP.....	70
DAFTAR PUSTAKA .....	72
LAMPIRAN.....	76
BIODATA PENULIS .....	78

## DAFTAR GAMBAR

Gambar 2. 1 Contoh decision tree data cuaca (Han et al., 2011).....	15
Gambar 2. 2 <i>Flowchart Decision Tree C4.5</i> .....	16
Gambar 2. 3 <i>Flowchart Naive Bayes</i> .....	17
Gambar 3. 1 Flowchart penelitian.....	21
Gambar 3. 2 <i>Flowchart SMOTE</i> .....	25
Gambar 3. 3 Contoh ilustrasi pohon keputusan .....	33
Gambar 4. 1 Histogram Dataset.....	40
Gambar 4. 2 Detail <i>value counts</i> label .....	40
Gambar 4. 3 Pie Chart dataset.....	41
Gambar 4. 4 Detail <i>value counts</i> label setelah dihapus.....	42
Gambar 4. 5 Setelah dilakukan <i>label encoder</i> .....	43
Gambar 4. 6 Hasil SMOTE .....	44
Gambar 4. 7 Visualisasi histogram after SMOTE .....	44
Gambar 4. 8 Hasil Binning.....	45
Gambar 4. 9 <i>Dataframe</i> kembali <i>imbalanced</i> .....	46
Gambar 4. 10 <i>Random undersampling</i> .....	46

## DAFTAR TABEL

Tabel 2. 1 Confusion Matrix .....	18
Tabel 2. 2 Kategori nilai AUC .....	20
Tabel 3. 1 Prosedur penelitian.....	22
Tabel 3. 2 Atribut penelitian .....	23
Tabel 3. 3 Contoh data .....	26
Tabel 3. 4 Contoh pembagian kategori ( <i>discretization</i> ).....	27
Tabel 3. 5 <i>Categorized</i> data .....	28
Tabel 3. 6 Data <i>entropy</i> .....	31
Tabel 3. 7 Data <i>information gain</i> .....	32
Tabel 3. 8 P(BB Lahir).....	34
Tabel 3. 9 P(Usia) .....	34
Tabel 3. 10 P(BB) .....	34
Tabel 3. 11 P(TB).....	35
Tabel 3. 12 P(LILA).....	35
Tabel 3. 13 P(JK) .....	35
Tabel 3. 14 Data uji.....	36
Tabel 3. 15 Data uji terkategori.....	36
Tabel 3. 16 Hasil prediksi perhitungan naive bayes .....	37
Tabel 3. 17 <i>Confusion matrix</i> hasil uji.....	37
Tabel 3. 18 Pembagian data latih dan data uji per skenario.....	38
Tabel 4. 1 Penjelasan Function pada model Naive Bayes .....	56
Tabel 4. 2 Catatan hasil pengujian model pada skenario 1 .....	62
Tabel 4. 3 Catatan hasil pengujian model pada skenario 2 .....	64
Tabel 4. 4 Catatan hasil pengujian model pada skenario 3 .....	66
Tabel 4. 5 Rekap hasil uji model pada setiap skenario (Decision Tree C4.5) .....	67
Tabel 4. 6 Rekap hasil uji model pada setiap skenario (Naïve Bayes) .....	67
Tabel 4. 7 Catatan hasil waktu komputasi (dalam detik).....	69