

SKRIPSI

**KOMPARASI ALGORITMA C4.5 DENGAN ALGORITMA
NAÏVE BAYES CLASSIFIER (NBC) UNTUK KLASIFIKASI
STATUS GIZI BALITA**



ANITA MARTILOFA

NIM : 225611009

**PROGRAM STUDI SISTEM INFORMASI
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA
2023**

HALAMAN PERSETUJUAN

USULAN SKRIPSI

**KOMPARASI ALGORITMA C4.5 DENGAN ALGORITMA
NAÏVE BAYES CLASSIFIER (NBC) UNTUK KLASIFIKASI
STATUS GIZI BALITA**

Diusulkan oleh
ANITA MARTILOFA
225611009

Telah disetujui
pada tanggal 7 Februari 2024.

Pembimbing



Sumiyatun, S.Kom., M. Cs.

NIP/NPP.091158

HALAMAN PERSETUJUAN
UJIAN SKRIPSI

Judul : **Komparasi Algoritma C4.5 dengan Algoritma Naïve Bayes Classifier (NBC) Untuk Klasifikasi Status Gizi Balita**

Nama : **Anita Martilofa**

NIM : **225611009**

Program Studi : **Sistem Informasi**

Program : **Sarjana**

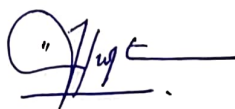
Semester : **Gasal**

Tahun Akademik : **2023/2024**

Telah diperiksa dan disetujui untuk diujikan di hadapan Dewan Penguji Skripsi

Yogyakarta, 7 Februari 2024

Dosen Pembimbing,



Sumiyatun, S.Kom., M. Cs.

NIP/NPP : 091158

HALAMAN PENGESAHAN



SKRIPSI

**KOMPARASI ALGORITMA C4.5 DENGAN ALGORITMA NAÏVE
BAYES CLASSIFIER (NBC) UNTUK KLASIFIKASI STATUS GIZI
BALITA**

Telah dipertahankan di depan Dewan Penguji Skripsi dan dinyatakan diterima untuk memenuhi sebagai persyaratan guna memperoleh Gelar

**Sarjana Komputer
Program Studi Sistem Informasi
Universitas Teknologi Digital Indonesia
Yogyakarta**

Yogyakarta, 7 Februari 2024

Dewan Penguji	NIP/NPP	Tanda Tangan
1. Danny Kriestanto, S.Kom., M.Eng (Ketua)	050206802	
2. Sumiyatun, S.Kom., M.Cs. (Sekretaris)	091158	

Mengetahui
Ketua Program Studi Sistem Informasi

Deborah Kurniawati, S.Kom., M.Cs.
NPP : 051149

19 FEB 2024

PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 7 Pebruari 2024



Anita Martilofa

NIM : 225611009

KATA PENGANTAR

Segala puji dan syukur Penulis panjatkan kehadirat Allah SWT, atas karunia-Nya kepada Penulis sehingga dapat menyelesaikan Skripsi yang berjudul “Komparasi *Algoritma C4.5* Dengan *Algoritma Naïve Bayes Classifier (NBC)* Untuk Klasifikasi Status Gizi Balita”. Shalawat serta salam Penulis panjatkan kepada junjungan kita Nabi Muhammad SAW yang kita harapkan safa’atnya di yaumul akhir.

Adapun penyusunan Skripsi ini diajukan untuk memenuhi tugas akhir sebagai salah satu syarat untuk memperoleh gelar sarjana.

Dalam penyusunan Skripsi ini Penulis banyak mengalami hambatan dan kesulitan akan tetapi atas bimbingan serta arahan dari para pembimbing dan dukungan semua pihak yang bersangkutan, Penulis dapat menyelesaikan Skripsi ini dengan baik. Untuk itu Penulis mengucapkan terima kasih kepada :

1. Ibu Deborah Kurniawati, S.Kom,M.Cs selaku Kaprodi Sistem Infomasi Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia di Yogyakarta
2. Ibu Sumiyatun, S.Kom,M.Cs selaku dosen Pembimbing
3. Seluruh dosen dan staf Program Studi Sistem Informasi Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia di Yogyakarta

4. Dinas Kesehatan Kabupaten Wonosobo yang telah mengizinkan Penulis melakukan studi kasus dan telah memberikan bimbinganya dalam penyusunan Skripsi ini.
5. Orang tua, Suami dan anak- anak yang telah memberikan dukungan, doa serta segala fasilitas dalam penyusunan Skripsi ini.
6. Teman – teman Prodi Sistem Informasi Kelas Karyawan angkatan 2021/2022 Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia di Yogyakarta
7. Seluruh pihak yang telah membantu dalam penyusunan Skripsi ini yang tidak bisa disebutkan satu persatu.

Penulis menyadari bahwa dalam penyusunan Skripsi ini masih jauh dari sempurna. Untuk itu Penulis mengharapkan kritik dan saran yang membangun untuk perbaikan Skripsi ini.

Akhir kata, Penulis berharap agar Skripsi ini dapat berguna bagi semua pihak serta menjadi sesuatu yang berarti dari usaha penulis selama ini.

Wonosobo, 8 Februari 2024



Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	<i>i</i>
HALAMAN PENGESAHAN	<i>iv</i>
PERNYATAAN KEASLIAN SKRIPSI	<i>v</i>
KATA PENGANTAR	<i>x</i>
ABSTRAK	<i>xii</i>
ABSTRACT	<i>xiii</i>
BAB 1 PENDAHULUAN	<i>1</i>
1.1. Latar Belakang Masalah	<i>1</i>
1.2. Rumusan Masalah	<i>3</i>
1.3. Ruang Lingkup	<i>3</i>
1.4. Tujuan Penelitian	<i>3</i>
1.5. Manfaat Penelitian	<i>3</i>
1.6. Sistematika Penulisan	<i>4</i>
BAB 2 TINJAUAN PUSTAKA DAN DASAR TEORI	<i>5</i>
2.1. Tinjauan Pustaka	<i>5</i>
2.2. Landasan Teori	<i>10</i>
2.2.1 Definisi <i>Machine Learning</i>	<i>10</i>
2.2.2 Proses <i>Machine Learning</i>	<i>10</i>
2.2.3 Teknik <i>Machine Learning</i>	<i>12</i>
2.2.4 <i>Decision Tree</i>	<i>15</i>
2.2.5 <i>Algoritma C4.5</i>	<i>16</i>
2.2.5.1 <i>Entropy</i>	<i>24</i>
2.2.5.2 <i>Information Gain</i>	<i>25</i>
2.2.6 <i>Algoritma Naïve Bayes Classifier (NBC)</i>	<i>25</i>
2.2.7 <i>RapidMiner</i>	<i>28</i>
2.2.7.1 <i>Sejarah RapidMiner</i>	<i>28</i>
2.2.7.2 <i>Pengertian RapidMiner</i>	<i>29</i>
2.2.7.3 <i>Terminologi dasar RapidMiner</i>	<i>30</i>
2.2.7.4 <i>Desain proses analisa dalam RapidMiner</i>	<i>31</i>
2.2.7.5 <i>Repositori Pertama</i>	<i>32</i>
2.2.7.6 <i>Perspektif dan View</i>	<i>32</i>
2.2.7.7 <i>Perspektif Desain</i>	<i>32</i>
2.2.7.8 <i>Operator dan Proses</i>	<i>33</i>
2.2.8 <i>Status Gizi</i>	<i>34</i>
2.2.8.1 <i>Pengertian Status Gizi</i>	<i>34</i>
2.2.8.2 <i>Penilaian Status Gizi</i>	<i>35</i>
2.2.8.3 <i>Indeks Antropometri</i>	<i>37</i>

BAB 3 METODOLOGI PENELITIAN.....	41
3.1 Tujuan Penelitian.....	41
3.2 Waktu dan Tempat Penelitian	41
3.3 Alat dan Bahan Penelitian	41
3.4 Objek Penelitian	42
3.5 Metode Pengumpulan Data.....	42
3.6 Metode Penelitian	43
3.6.1 Pemahaman Bisnis (<i>Business Understanding</i>).....	43
3.6.2 Pemahaman Data (<i>Data Understanding</i>).....	43
3.6.3 Pengolahan Data (<i>Data Preparation</i>)	44
3.6.4 Pemodelan (<i>Modelling</i>).....	45
3.6.5 Evaluasi (<i>Evaluation</i>)	46
3.6.6 Deployment.....	46
3.7 Jadwal Penelitian	47
BAB 4 IMPLEMENTASI DAN PEMBAHASAN.....	48
4.1 Implementasi Data	48
4.2 Pengumpulan Data	48
4.2.1 Cleaning Data.....	49
4.2.2 Seleksi Data	49
4.2.3 Penyiapan File.....	50
4.2.4 Classification.....	51
4.2.5 Penerapan Algoritma Decision Tree C4.5.....	51
4.2.6 Penerapan Algoritma <i>Naïve Bayes Classifier (NBC)</i>	59
4.2.7 Pengolahan Data Menggunakan Aplikasi RapidMiner 9.10	65
BAB 5 SIMPULAN DAN SARAN	72
5.1 SIMPULAN	72
5.2 SARAN	73
DAFTAR PUSTAKA	74
Lampiran.....	76

DAFTAR GAMBAR

<i>Gambar 2. 1 Skema Artificial Intelligence dan Machine Learning</i>	12
<i>Gambar 2. 2 Pohon Keputusan Hasil Perhitungan Node 1</i>	21
<i>Gambar 2. 3 Pohon Keputusan Hasil Perhitungan Node 1.1</i>	23
<i>Gambar 2. 4 Pohon Keputusan Hasil Perhitungan Node 1.1.2</i>	24
<i>Gambar 4. 1 Pohon Keputusan Node 1</i>	56
<i>Gambar 4. 2 Pohon Keputusan Node 1.1</i>	59
<i>Gambar 4. 3 Tampilan Utama aplikasi RapidMiner 9.10</i>	65
<i>Gambar 4. 4 Tampilan main process</i>	65
<i>Gambar 4. 5 Tampilan import file</i>	66
<i>Gambar 4. 6 Tampilan hasil import file</i>	66
<i>Gambar 4. 7 Tampilan pemilihan label pada data</i>	67
<i>Gambar 4. 8 Tampilan hasil akhir import data</i>	67
<i>Gambar 4. 9 Tampilan data statistik</i>	68
<i>Gambar 4. 10 Tampilan process tahap 1</i>	68
<i>Gambar 4. 11 Tampilan process tahap 2</i>	69
<i>Gambar 4. 12 Tampilan process tahap 3</i>	69
<i>Gambar 4. 13 Tampilan process tahap 4</i>	70
<i>Gambar 4. 14 Tampilan process tahap 4</i>	70
<i>Gambar 4. 15 Tampilan hasil performa algoritma C.45</i>	71
<i>Gambar 4. 16 Tampilan hasil performa algoritma naïve bayes</i>	71

DAFTAR TABEL

<i>Tabel 2. 1 Studi Sebelumnya</i>	8
<i>Tabel 2. 2 Keputusan Bermain Tennis</i>	19
<i>Tabel 2. 3 Perhitungan Node 1</i>	20
<i>Tabel 2. 4 Perhitungan Node 1.1</i>	22
<i>Tabel 2. 5 Perhitungan Node 1.1.2</i>	23
<i>Tabel 2. 6 Contoh Data Set Naïve Bayes</i>	27
<i>Tabel 2. 7 Tabel Perhitungan Contoh Kasus Naïve Bayes</i>	27
<i>Tabel 2. 8 Tabel Hasil Contoh Kasus Naïve Bayes</i>	28
<i>Tabel 3. 1 Jadwal Penelitian</i>	47
<i>Tabel 4. 1 Data Status Gizi Balita (BB/TB)</i>	48
<i>Tabel 4. 2 Data Status Gizi setelah melalui tahap cleaning data dan selection data</i>	50
<i>Tabel 4. 3 Perhitungan Jumlah Kasus Total BB/TB</i>	54
<i>Tabel 4. 4 Perhitungan Entropy dan Gain Node 1 BB/TB</i>	55
<i>Tabel 4. 5 Perhitungan Entropy dan Gain Node 1.1</i>	57
<i>Tabel 4. 6 Perhitungan Entropy dan Gain Node 1.1</i>	57
<i>Tabel 4. 7 Perhitungan Entropy dan Gain Node 1.1</i>	58
<i>Tabel 4. 8 Perhitungan Entropy dan Gain Node 1.1</i>	58
<i>Tabel 4. 9 Contoh Data Testing Algoritma NBC</i>	59
<i>Tabel 4. 10 Hasil Data Testing Algoritma NBC</i>	64

KATA PENGANTAR

Segala puji dan syukur Penulis panjatkan kehadirat Allah SWT, atas karunia-Nya kepada Penulis sehingga dapat menyelesaikan Tugas Akhir yang berjudul “Komparasi *Algoritma C4.5* Dengan *Algoritma Naïve Bayes Classifier (NBC)* Untuk Klasifikasi Status Gizi Balita”. Shalawat serta salam Penulis panjatkan kepada junjungan kita Nabi Muhammad SAW yang kita harapkan safa’atnya di yaumul akhir.

Adapun penyusunan Tugas Akhir ini diajukan untuk memenuhi tugas akhir sebagai salah satu syarat untuk memperoleh gelar sarjana.

Dalam penyusunan Tugas Akhir ini Penulis banyak mengalami hambatan dan kesulitan akan tetapi atas bimbingan serta arahan dari para pembimbing dan dukungan semua pihak yang bersangkutan, Penulis dapat menyelesaikan Tugas Akhir ini dengan baik. Untuk itu Penulis mengucapkan terima kasih kepada :

1. Ibu Deborah Kurniawati, S.Kom,M.Cs selaku Kaprodi Sistem Infomasi Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia di Yogyakarta
2. Ibu Sumiyatu, S.Kom,M.Cs selaku dosen Pembimbing
3. Seluruh dosen dan staf Program Studi Sistem Informasi Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia di Yogyakarta

4. Dinas Kesehatan Kabupaten Wonosobo yang telah mengizinkan Penulis melakukan studi kasus dan telah memberikan bimbinganya dalam penyusunan Tugas Akhir ini.
5. Suami, anak- anak dan orang tua yang telah memberikan dukungan, doa serta segala fasilitas dalam penyusunan Tugas Akhir ini.
6. Teman – teman Prodi Sistem Informasi Kelas Karyawan angkatan 2021/2022 Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia di Yogyakarta
7. Seluruh pihak yang telah membantu dalam penyusunan Tugas Akhir ini yang tidak bisa disebutkan satu persatu.

Penulis menyadari bahwa dalam penyusunan Tugas Akhir ini masih jauh dari sempurna. Untuk itu Penulis mengharapkan kritik dan saran yang membangun untuk perbaikan Tugas Akhir ini.

Akhir kata, Penulis berharap agar Tugas Akhir ini dapat berguna bagi semua pihak serta menjadi sesuatu yang berarti dari usaha penulis selama ini.

Wonosobo, 8 Februari 2024

Penulis

ABSTRAK

Penelitian ini membahas perbandingan dua metode klasifikasi, yaitu metode algoritma C4.5 dan algoritma Naive Bayes, dalam memprediksi status gizi balita. Metode C4.5 merupakan algoritma pohon keputusan (decision tree) yang memanfaatkan aturan-aturan untuk membagi dataset menjadi kelompok yang lebih kecil. Sementara itu, metode Naive Bayes berdasarkan pada teorema Bayes dan mengasumsikan independensi antar- fitur dalam pengklasifikasian.

Penelitian ini menggunakan dataset yang mencakup variabel-variabel yang relevan dengan status gizi balita, seperti berat badan, tinggi badan, umur balita, dan faktor-faktor lainnya. Pengujian dilakukan dengan membandingkan performa kedua metode dalam hal akurasi, presisi dan recall kelas (tingkat pengenalan kelas) terhadap klasifikasi dan kehandalan prediksi.

Hasil penelitian menunjukkan perbandingan kelebihan dan kelemahan dari masing-masing metode. Analisis statistik digunakan untuk menilai signifikansi perbedaan antara dua metode tersebut. Temuan penelitian ini diharapkan dapat memberikan wawasan lebih lanjut terkait efektivitas dan aplikabilitas metode C4.5 dan Naive Bayes dalam konteks klasifikasi status gizi balita. Implikasi dari perbandingan ini dapat menjadi landasan bagi pengembangan model klasifikasi yang lebih baik guna meningkatkan akurasi prediksi status gizi balita di masa depan.

Kata kunci : Algoritma C4.5, Algoritma *Naive Bayes Classification*, *Decision Tree*, Klasifikasi, Status gizi balita.

ABSTRACT

This research explores the comparison of two classification methods, namely the C4.5 algorithm method and Naive Bayes algorithm, in predicting the nutritional status of toddlers. The C4.5 method is a decision tree algorithm that utilizes rules to divide the dataset into smaller groups. Meanwhile, the Naive Bayes method is based on Bayes' theorem and assumes independence among features in classification.

The study utilizes a dataset that includes relevant variables related to toddlers' nutritional status, such as weight, height, age, and other factors. Testing is conducted by comparing the performance of both methods in terms of accuracy, precision, and class recall (class recognition rate) for classification and prediction reliability.

The research results reveal the strengths and weaknesses of each method. Statistical analysis is employed to assess the significance of differences between the two methods. The findings of this research are expected to provide further insights into the effectiveness and applicability of the C4.5 and Naive Bayes methods in the context of classifying the nutritional status of toddlers. The implications of this comparison can serve as a foundation for the development of improved classification models to enhance the accuracy of predicting toddlers' nutritional status in the future.

Keywords : Classification, C4.5 Algorithm, Decision Tree, Naïve Bayes Classification Algorithm, Nutritional Status.