

**PERBANDINGAN AKURASI ALGORITMA DECISION TREE
C4.5 DAN NAIVE BAYES DALAM MENENTUKAN JURUSAN
DI MA MUALLIMIN MUHAMMADIYAH YOGYAKARTA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

DEWANGGA MANTARA SAKTI

19.11.3116

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2022

**PERBANDINGAN AKURASI ALGORITMA DECISION TREE C4.5
DAN NAIVE BAYES DALAM MENENTUKAN JURUSAN DI MA
MUALLIMIN MUHAMMADIYAH YOGYAKARTA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

DEWANGGA MANTARA SAKTI

19.11.3116

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA**

YOGYAKARTA

2023

HALAMAN PERSETUJUAN

SKRIPSI

**PERBANDINGAN AKURASI ALGORITMA DECISION TREE C4.5 DAN
NAIVE BAYES DALAM MENENTUKAN JURUSAN DI MA MUALLIMIN
MUHAMMADIYAH YOGYAKARTA**

yang disusun dan diajukan oleh

Dewangga Mantara Sakti

19.11.3116

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 10 Agustus 2023

Dosen Pembimbing,



Windha Mega Pradnya D. M.Kom
NIK. 190302185

HALAMAN PENGESAHAN

SKRIPSI

PERBANDINGAN AKURASI ALGORITMA DECISION TREE C4.5 DAN
NAIVE BAYES DALAM MENENTUKAN JURUSAN DI MA MUALLIMIN
MUHAMMADIYAH YOGYAKARTA

yang disusun dan diajukan oleh

Dewangga Mantara Sakti

19.11.3116

Telah dipertahankan di depan Dewan Penguji
pada tanggal 10 Agustus 2023

Susunan Dewan Penguji

Nama Penguji

Arif Dwi Laksito, M.Kom
NIK. 190302150

Nuri Cahyono, M.Kom
NIK. 190302278

Windha Mega Pradnya D, M.Kom
NIK. 190302185

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 10 Agustus 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Dewangga Mantara Sakti
NIM : 19.11.3116

Menyatakan bahwa Skripsi dengan judul berikut:

PERBANDINGAN AKURASI ALGORITMA DECISION TREE C4.5 DAN NAIVE BAYES DALAM MENENTUKAN JURUSAN DI MA MUALLIMIN MUHAMMADIYAH YOGYAKARTA

Dosen Pembimbing : Windha Mega Pradnya D, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 10 Agustus 2023

Yang Menyatakan,

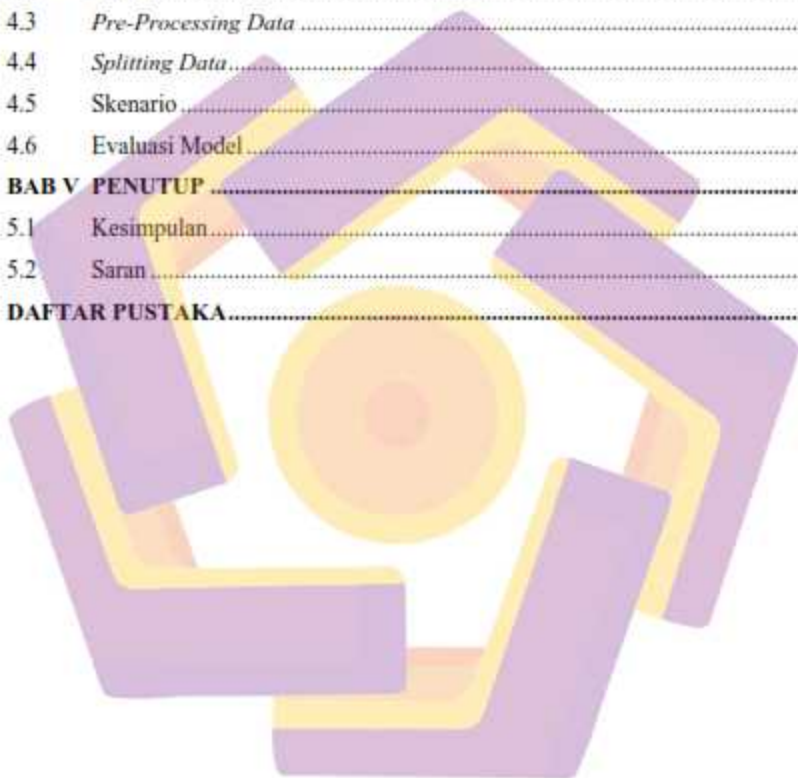


Dewangga Mantara Sakti

DAFTAR ISI

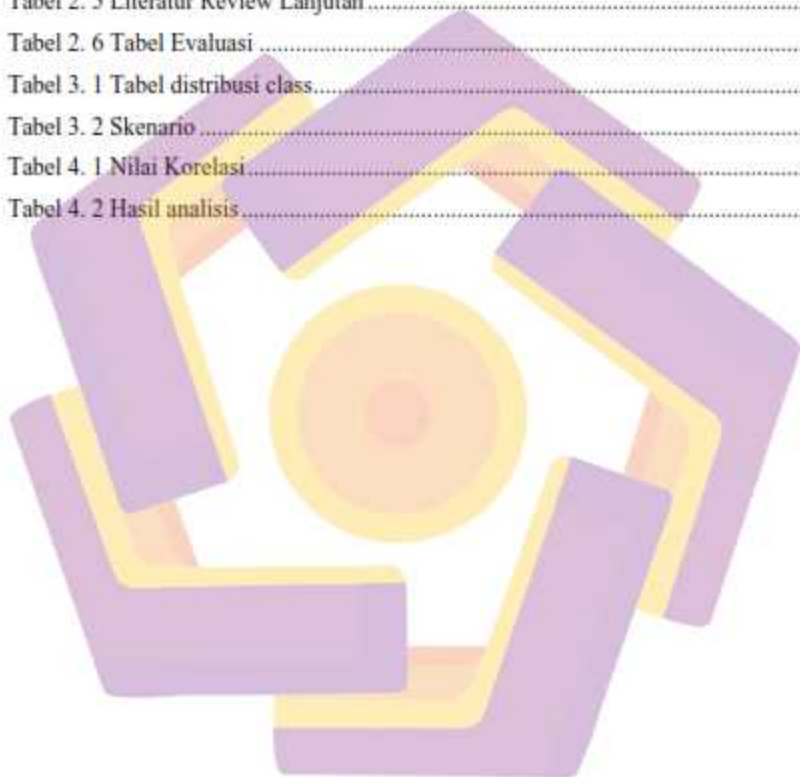
HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
DAFTAR ISI	v
DAFTAR TABEL	vii
DAFTAR GAMBAR	viii
INTISARI	ix
ABSTRACT	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	2
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	4
2.1 Studi Literatur	4
2.2 Dasar Teori	12
2.2.1 Data Mining	12
2.2.2 Algoritma	15
2.2.3 C4.5	15
2.2.4 Naïve Bayes	16
2.2.5 Python	18
BAB III METODE PENELITIAN	19
3.1 Objek Penelitian	19
3.2 Alur Penelitian	20
3.2.1 Identifikasi Masalah	20
3.2.2 Dataset	21
3.2.3 Analisa	21
3.2.4 <i>Pre-Processing Data</i>	22

3.2.5	<i>Splitting Data Training dan Testing</i>	22
3.2.6	Skenario Pengujian.....	22
3.2.7	Evaluasi Model	23
3.3	Spesifikasi Alat.....	23
BAB IV HASIL DAN PEMBAHASAN		24
4.1	Dataset	24
4.2	Analisa	24
4.3	<i>Pre-Processing Data</i>	26
4.4	<i>Splitting Data</i>	30
4.5	Skenario	31
4.6	Evaluasi Model.....	32
BAB V PENUTUP		37
5.1	Kesimpulan.....	37
5.2	Saran	37
DAFTAR PUSTAKA.....		38



DAFTAR TABEL

Tabel 2. 1 Literatur Review	7
Tabel 2. 2 Literatur Review Lanjutan	8
Tabel 2. 3 Literatur Review Lanjutan	9
Tabel 2. 4 Literatur Review Lanjutan	10
Tabel 2. 5 Literatur Review Lanjutan	11
Tabel 2. 6 Tabel Evaluasi	18
Tabel 3. 1 Tabel distribusi class	21
Tabel 3. 2 Skenario	23
Tabel 4. 1 Nilai Korelasi	24
Tabel 4. 2 Hasil analisis	36



DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	20
Gambar 3. 2 Dataset Siswa	21
Gambar 4. 1 DropNA.....	26
Gambar 4. 2 Drop Duplicate	27
Gambar 4. 3 Import Label Encoder	27
Gambar 4. 4 Data setelah di ubah	28
Gambar 4. 5 Data sebelum di ubah	28
Gambar 4. 6 <i>import library</i> pandas.....	29
Gambar 4. 7 mengkategorikan nilai.....	29
Gambar 4. 8 merubah tipe data nilai.....	30
Gambar 4. 9 <i>splitting</i> data algoritma C4.5.....	31
Gambar 4. 10 <i>splitting</i> data algoritma Naïve Bayes	31
Gambar 4. 11 Enam Atribut 20 Kategorial Naïve Bayes	32
Gambar 4. 12 Enam Atribut 20 Kategorial C4.5	32
Gambar 4. 13 Enam Atribut 4 Kategorial Naïve Bayes	33
Gambar 4. 14 Enam Atribut 4 Kategorial C4.5	33
Gambar 4. 15 Delapan Atribut 4 Kategorial Naïve Bayes	34
Gambar 4. 16 Delapan Atribut 4 Kategorial C4.5	34
Gambar 4. 17 Delapan Atribut 20 Kategorial Naïve Bayes	35
Gambar 4. 18 Delapan Atribut 20 Kategorial C4.5	35

INTISARI

Pada studi kasus yang dilakukan di Madrasah Muallimin Muhammadiyah Yogyakarta penentuan jurusan untuk siswa MTS masih menggunakan metode manual kurang efektif. Masalah yang sering muncul saat pemilihan jurusan adalah waktu yang di butuhkan lebih lama dan juga sumber daya yang lebih banyak. Setelah berdiskusi dengan guru bimbingan konseling pada objek penelitian, telah dibahas bagaimana cara melakukan pemilihan jurusan agar lebih efektif dengan menggunakan sebuah metode yang baik dalam menentukannya. Oleh karena itu dibutuhkan algoritma yang cocok untuk menentukan jurusan siswa MTS Muallimin Muhammadiyah Yogyakarta. Penelitian ini memiliki tujuan untuk membandingkan hasil akurasi terbaik dari dua buah algoritma yaitu algoritma C4.5 dan algoritma Naïve Bayes. Data yang sudah di berikan oleh objek penelitian sudah cukup untuk menentukan jurusan siswa Madrasah Muallimin Muhammadiyah Yogyakarta. Dengan beberapa atribut tambahan selain data nilai yaitu data tes potensi akademik dan data rekomendasi setiap orang tua siswa. Metode untuk analisis menggunakan metode classification report. Hasil dari penelitian ini merupakan hasil akurasi yang sudah dicoba dengan beberapa skenario pengujian dengan beberapa atribut dan kategorial nilai yang berbeda.

Kata kunci: algoritma C4.5, algoritma Naïve Bayes, classification report, Muallimin Muhammadiyah Yogyakarta.

ABSTRACT

In the case study conducted at Madrasah Muallimin Muhammadiyah Yogyakarta, the determination of majors for MTS students still using manual methods is less effective. The problem that often arises when choosing a major is the time it takes longer and more resources. After discussing with the counselling guidance teacher on the object of research, it has been discussed how to make the selection of majors more effective by using a good method in determining it. Therefore, a suitable algorithm is needed to determine the major of MTS Muallimin Muhammadiyah Yogyakarta students. This study aims to compare the best accuracy results of two algorithms, namely the C4.5 algorithm and the Naïve Bayes algorithm. The data that has been provided by the object of research is sufficient to determine the major of students of Madrasah Muallimin Muhammadiyah Yogyakarta. With some additional attributes besides grade data, namely academic potential test data and recommendation data for each student's parents. The method for analysis uses the classification report method. The results of this study are the accuracy results that have been tried with several test scenarios with several different attributes and value categories.

Keywords: C4.5 algorithm, Naïve Bayes algorithm, classification report, Muallimin Muhammadiyah Yogyakarta.