

**PERBANDINGAN ALGORITMA NAÏVE BAYES DAN C4.5  
UNTUK PREDIKSI KELULUSAN MAHASISWA STUDI  
KASUS: STABN SRIWIJAYA**



**UMN**  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

**SKRIPSI**

**Agga Pagunnata Panuntun**

**00000025763**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA NUSANTARA  
TANGERANG  
2023**

**PERBANDINGAN ALGORITMA NAÏVE BAYES DAN C4.5  
UNTUK PREDIKSI KELULUSAN MAHASISWA STUDI  
KASUS: STABN SRIWIJAYA**



**UMN**  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

**SKRIPSI**

Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer (S.Kom)

**Agga Pagunnata Panuntun**

**0000025763**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA NUSANTARA  
TANGERANG**

**2023**

**HALAMAN PERNYATAAN TIDAK PLAGIAT**

Dengan ini saya,

Nama : Agga Pagunnata Panuntun

Nomor Induk Mahasiswa : 00000025763

Program studi : Sistem Informasi

Laporan Skripsi dengan judul:

Perbandingan Algoritma Naïve Bayes Dan C4.5 Untuk Prediksi Kelulusan Mahasiswa Studi Kasus: STABN Sriwijaya merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas Akhir yang telah saya tempuh.

Tangerang, 9 Oktober 2023

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

Agga Pagunnata Panuntun



i

Perbandingan Algoritma Naïve Bayes Dan C4.5 Untuk Prediksi Kelulusan  
Mahasiswa Studi Kasus: STABN Sriwijaya, Agga Pagunnata Panuntun,  
Universitas Multimedia Nusantara

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul  
Perbandingan Algoritma Naive Bayes Dan C4.5 Untuk Prediksi Kelulusan  
Mahasiswa Studi Kasus: STABN Sriwijaya

Oleh  
Nama : Agga Pagunnata Panuntun  
NIM : 00000025763  
Program Studi : Sistem Informasi  
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Selasa, 17 Oktober 2023  
Pukul 15.00 s.d 16.00 dan dinyatakan  
LULUS  
Dengan susunan penguji sebagai berikut.

Ketua Sidang

Penguji

Wella, S.Kom., M.MSI. (COBIT5)<sup>19</sup>

Monika Evelyn Johan, S.Kom., M.M.S.I.

Pembimbing

Ririn Ikana Desanti, S.Kom., M.Kom.

Ketua Program Studi Sistem Informasi

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

Perbandingan Algoritma Naive..., Agga Pagunnata Panuntun, Universitas Multimedia Nusantara

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## KATA PENGANTAR

Puji Syukur atas selesainya Penelitian Laporan Skripsi ini dengan judul: “Perbandingan Algoritma Naïve Bayes Dan C4.5 Untuk Prediksi Kelulusan Mahasiswa Studi Kasus: STABN Sriwijaya” dilakukan untuk memenuhi syarat untuk mendapatkan gelar Sarjana Komputer S.Kom Jurusan Sistem Informasi Pada Teknik dan Informatika Universitas Multimedia Nusantara.

Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan laporan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan laporan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, M.A., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga Peneliti dapat menyelesaikan tugas akhir ini.
6. Teman serta para sahabat yang telah memberikan dukungan dan semangat sehingga Peneliti dapat menyelesaikan laporan Skripsi ini.

Semoga laporan skripsi yang telah Peneliti buat dapat memberikan manfaat yang berharga bagi para pembaca.

Tangerang, 9 Oktober 2023

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA



Agga Pagunnata Panuntun

**PEMILIHAN ALGORITMA DATA MINING TERBAIK UNTUK  
PREDIKSI KELULUSAN MAHASISWA STUDI KASUS STABN**

**SRIWIJAYA**

Agga Pagunnata Panuntun

**ABSTRAK**

Sekolah Tinggi merupakan lembaga pendidikan tinggi yang berperan penting dalam memberikan Pendidikan. Persentase siswa yang lulus tepat waktu adalah hal yang penting dalam menentukan kualitas sebuah sekolah tinggi. Akan tetapi fakta dilapangan, masih banyak mahasiswa yang lulusnya terlambat yang disebabkan karena beberapa faktor seperti nilai Indeks Prestasi Semester (IPK) dan Indeks Prestasi Semester (IPS) yang rendah.

Tujuan dari penelitian ini untuk melakukan analisis perbandingan dalam prediksi kelulusan mahasiswa yang tepat waktu dan yang terlambat di STABN Sriwijaya. Adapun metodologi yang digunakan pada penelitian ini yaitu dengan menggunakan CRISP-DM yang dimulai dari business understanding, data understanding, data preparation, modeling, evaluation dan deployment sedangkan pada proses pengolahan data menggunakan pendekatan data mining dengan metode *Naïve Bayes* dan C4.5 yang didukung oleh *tools RapidMiner*.

Hasil dari penelitian ini adalah dari *accuracy*, *precision* dan *recall*. Hasil dari *Naïve Bayes* yaitu *accuracy* sebesar 88.10%, *precision* sebesar 58.62% dan *recall* sebesar 81.73%. Untuk C4.5 yaitu *accuracy* sebesar 96.08%, *precision* sebesar 94.32% dan *recall* sebesar 79.81%. Hasil yang sudah didapat dari perbandingan tersebut dapat disimpulkan jika C4.5 memiliki hasil yang paling optimal dibandingkan dengan metode *Naïve Bayes*.

**Kata kunci:** C4.5, Data Mining, Kelulusan Mahasiswa, *Naïve Bayes*, RapidMiner

UMMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

# SELECTION OF THE BEST DATA MINING ALGORITHM FOR PREDICTING STUDENT GRADUATION CASE STUDY STABN

SRIWIJAYA

Agga Pagunnata Panuntun

## ***ABSTRACT (English)***

*High School is a higher education institution that plays an important role in providing education. The percentage of students who graduate on time is important in determining the quality of a high school. However, the facts on the ground are that there are still many students who graduate late due to several factors such as low Semester Achievement Index (GPA) and Semester Achievement Index (IPS).*

*The aim of this research is to carry out a comparative analysis in predicting student graduation on time and late at STABN Sriwijaya. The methodology used in this research is using CRISP-DM starting from business understanding, data understanding, data preparation, modeling, evaluation and deployment while the data processing process uses a data mining approach with the Naïve Bayes and C4.5 methods which are supported by RapidMiner tools*

*The results of this research are accuracy, precision and recall. The results of Naïve Bayes are accuracy of 88.10%, precision of 58.62% and recall of 81.73%. For C4.5, the accuracy is 96.08%, precision is 94.32% and recall is 79.81%. The results obtained from this comparison can be concluded that C4.5 has the most optimal results compared to the Naïve Bayes method.*

**Keywords:** *C4.5, Data Mining, Naïve Bayes, RapidMiner, Student Graduation*

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A

## DAFTAR ISI

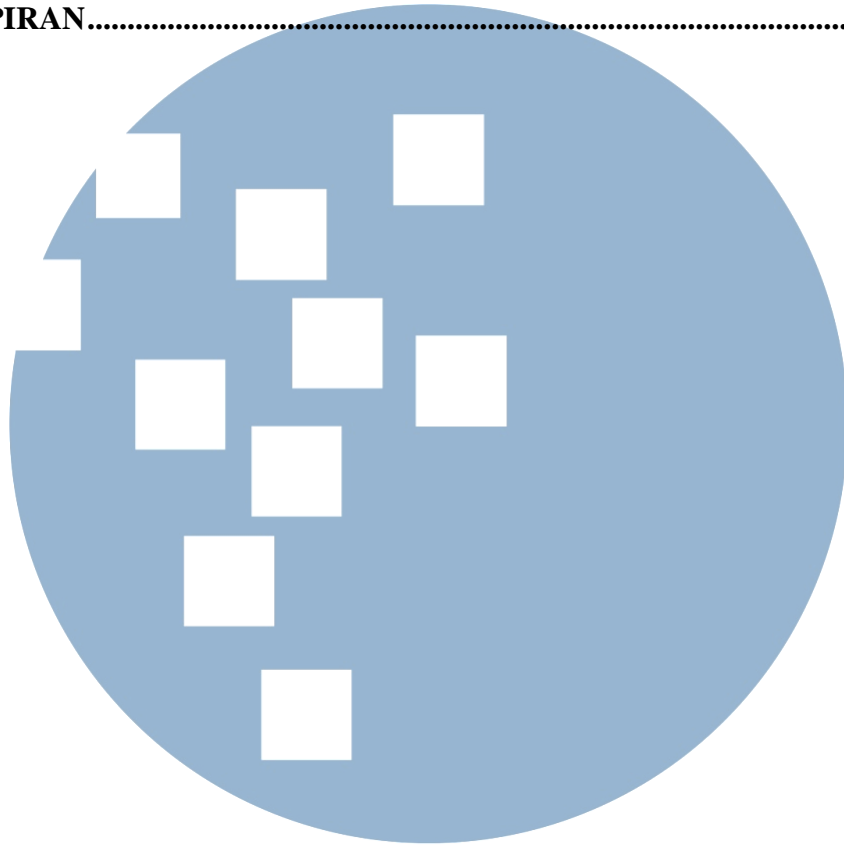
### Contents

<b>PERBANDINGAN ALGORITMA NAÏVE BAYES DAN C4.5 UNTUK PREDIKSI KELULUSAN MAHASISWA STUDI KASUS: STABN SRIWIJAYA</b> .....	1
.....	i
<b>HALAMAN PERNYATAAN TIDAK PLAGIAT</b> .....	i
<b>HALAMAN PERSETUJUAN</b> .....	ii
.....	i
<b>HALAMAN PENGESAHAN</b> .....	i
<b>HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS</b> .....	ii
<b>KATA PENGANTAR</b> .....	iii
<b>ABSTRAK</b> .....	iv
<i>ABSTRACT (English)</i> .....	v
<b>DAFTAR TABEL</b> .....	ix
<b>DAFTAR GAMBAR</b> .....	x
<b>DAFTAR LAMPIRAN</b> .....	xi
<b>1.1 Latar Belakang</b> .....	1
<b>1.2 Rumusan Masalah</b> .....	4
<b>1.3 Batasan Masalah</b> .....	4
<b>1.4 Tujuan dan Manfaat Penelitian</b> .....	4
<b>1.4.1 Tujuan Penelitian</b> .....	4
<b>1.4.2 Manfaat Penelitian</b> .....	4
<b>1.5 Sistematika Penelitian</b> .....	5
<b>BAB II LANDASAN TEORI</b> .....	6
<b>2.1 Kelulusan Mahasiswa pada Sekolah Tinggi</b> .....	6
<b>2.2 Data Mining</b> .....	6
<b>2.2.1 Proses mining</b> .....	8
<b>2.2.2 Teknik Data Mining</b> .....	9
<b>2.2.3 CRISP-DM</b> .....	10



2.2.4 Metode Naïve Bayes .....	10
2.2.5 Metode C4.5 .....	11
2.2.6 Model Evaluasi .....	12
2.2.7 Confusion Matrix .....	12
2.3 RapidMiner.....	14
2.4 Penelitian Terdahulu .....	14
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>28</b>
3.1 Gambaran Umum Objek Penelitian.....	28
3.1.1 Profile Kampus STABN Sriwijaya.....	29
3.1.2 Visi dan Misi Kampus STABN Sriwijaya.....	30
3.2 Metode Penelitian.....	31
3.2.1 Alur Penelitian.....	34
3.4 Teknik Pengumpulan Data .....	37
3.5 Teknik Analisis Data.....	37
3.5.1 <i>Business Understanding</i> .....	37
3.5.2 <i>Data Understanding</i> .....	37
3.5.3 <i>Data Preparation</i> .....	38
3.5.4 <i>Modeling</i> .....	38
3.5.5 <i>Evaluation</i> .....	38
3.5.6 <i>Deployment</i> .....	38
<b>BAB IV ANALISIS DAN HASIL PENELITIAN .....</b>	<b>39</b>
4.1 Business Understanding .....	39
4.2 Data Understanding .....	39
4.3 Data Preparation.....	41
4.4 Modeling .....	43
4.5 Evaluation .....	47
4.6 Deployment .....	57
4.7 Analisis dan Diskusi .....	57
<b>BAB V .....</b>	<b>60</b>
<b>KESIMPULAN DAN SARAN .....</b>	<b>60</b>
5.1 Kesimpulan.....	60
5.2 Saran .....	60

**DAFTAR PUSTAKA** ..... 61  
**LAMPIRAN**..... 65



**UMN**

**UNIVERSITAS  
MULTIMEDIA  
NUSANTARA**

## DAFTAR TABEL

Tabel 2. 1 Confusion Matrix .....	13
Tabel 2. 2 Penelitian Terdahulu .....	14
Tabel 3. 1 Perbandingan Dua Metode.....	31
Tabel 3. 2 Hasil Tools Weka.....	32
Tabel 4. 1 Pembagian Variabel Prediktor dan Variabel Target .....	41
Tabel 4. 2 Perhitungan Manual .....	50
Tabel 4. 3 Perhitungan Manual .....	55
Tabel 4. 4 Komparasi Metode C4.5 dan Naïve Bayes .....	57
Tabel 4. 5 Hasil Perbandingan Akurasi.....	58

# UMMN

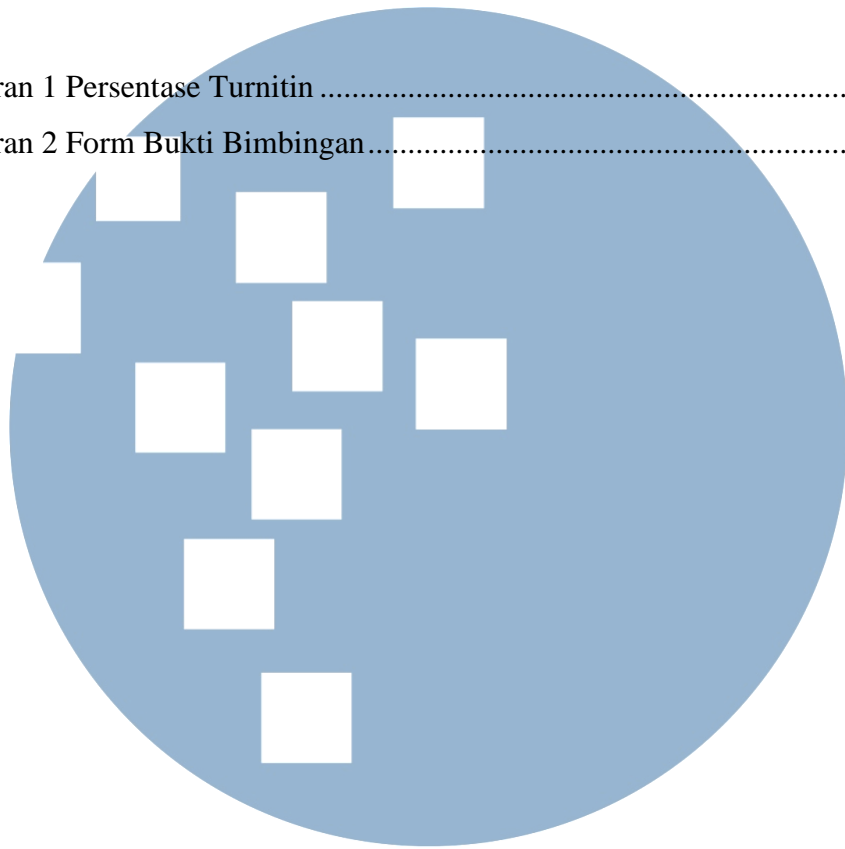
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## DAFTAR GAMBAR

Gambar 1. 1 Grafik Kelulusan Mahasiswa .....	2
Gambar 2. 1 Tahapan Knowledge Discovery in Database .....	7
Gambar 2. 2 Tahapan Proses CRISP-DM.....	10
Gambar 3. 1 Data Latih .....	28
Gambar 3. 2 Data Testing .....	29
Gambar 3. 3 Alur Penelitian.....	34
Gambar 3. 4 CRISP-DM .....	37
Gambar 4. 1 Data Latih.....	40
Gambar 4. 2 Data Testing .....	41
Gambar 4. 3 Data Preparation.....	42
Gambar 4. 4 Hasil Select Atribut .....	43
Gambar 4. 5 Operator Missing Value .....	43
Gambar 4. 6 Filter Example.....	44
Gambar 4. 7 Hasil Missing data.....	44
Gambar 4. 8 Proses Modeling Metode C4.5 .....	45
Gambar 4. 9 Proses Modeling Metode Naïve Bayes .....	45
Gambar 4. 10 Modeling Hitung Performance Metode C4.5.....	46
Gambar 4. 11 Modeling Hitung Performance Metode Naïve Bayes .....	46
Gambar 4. 12 Hasil Akurasi Metode C4.5.....	47
Gambar 4. 13 Hasil Precision Metode C4.5.....	48
Gambar 4. 14 Hasil Recall Metode C4.5 .....	48
Gambar 4. 15 Pohon Keputusan.....	49
Gambar 4. 16 Hasil Tree .....	49
Gambar 4. 17 Hasil Akurasi Metode Naïve Bayes .....	52
Gambar 4. 18 Hasil Precision Metode Naïve Bayes .....	53
Gambar 4. 19 Hasil Recall Metode Naïve Bayes.....	53
Gambar 4. 20 Hasil Metode Naïve Bayes.....	54
Gambar 4. 21 Hasil Sampel Distribusi Metode Naïve Bayes .....	54

## DAFTAR LAMPIRAN

Lampiran 1 Persentase Turnitin .....	65
Lampiran 2 Form Bukti Bimbingan.....	74



# UMMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA