

**IMPLEMENTASI ALGORITMA NAIVE BAYES DALAM MEMPREDIKSI  
KETERLAMBATAN BIAYA KULIAH DI UNIVERSITAS  
MUHAMMADIYAH KALIMANTAN TIMUR**

**SKRIPSI**

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar  
Sarjana Komputer

**DISUSUN OLEH :**

**MUHAMMAD NORHALIMI**

**1811102441064**



**PROGRAM STUDI S1 TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR  
SAMARINDA**

**2022**

**Implementasi Algoritma *Naïve Bayes* dalam Memprediksi  
Keterlambatan Biaya Kuliah di Universitas Muhammadiyah  
Kalimantan Timur**

**SKRIPSI**

Diajukan untuk memenuhi sebagian persyaratan pencapaian gelar  
Sarjana Komputer

**Disusun Oleh :**

**Muhammad Norhalimi**

**1811102441064**



**PROGRAM STUDI S1 TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR  
SAMARINDA  
2022**

# HALAMAN PENGESAHAN

## IMPLEMENTASI ALGORITMA *NAÏVE BAYES* DALAM MEMPREDIKSI KETERLAMBATAN BIAYA KULIAH DI UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR

DISUSUN OLEH :

MUHAMMAD NORHALIMI

1811102441064

Telah melaksanakan ujian skripsi dan dinyatakan lulus,

Pada tanggal 2 Juli 2022

Dosen Pembimbing



Taghfirul Azhima Yoga Siswa, S.Kom., M.Kom  
NIDN. 1118038805

Penguji



Rofilde Hasudungan, S.Kom., M.Sc  
NIDN. 1107048601

Dekan



Prof. Dr. Sarjito, MT., Ph.D  
NIDN. 0610116204

Ketua Program Studi



Asslia Johar Latipah, S.Kom., M.Cs  
NIDN. 1124098902

## PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Norhalimi  
NIM : 1811102441064  
Konsentrasi : Sistem Informasi

Menyatakan bahwa Skripsi dengan judul berikut:

**Implementasi Algoritma *Naïve Bayes* dalam Memprediksi Keterlambatan Biaya Kuliah di Universitas Muhammadiyah Kalimantan Timur.**

Dosen Pembimbing : Taghfirul Azhima Yoga Siswa, S.Kom., M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar Sarjana Komputer, baik di Universitas Muhammadiyah Kalimantan Timur(UMKT) 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 Muhammadiyah Kalimantan Timur (UMKT)
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 Muhammadiyah.

Samarinda, 2 Juli 2022

Yang Menyatakan,



Muhammad Norhalimi

1811102441064

## PRAKATA

Penulis mengucapkan *Alhamdulillah Robbil'alamin* kepada Allah Swt. Karena telah diberikan kesehatan, tenaga, pikiran, dan waktu luang sehingga dapat menyelesaikan penulisan skripsi ini dengan judul Implementasi algoritma *Naive Bayes* dalam memprediksi Keterlambatan Biaya Kuliah di Universitas Muhammadiyah Kalimantan Timur. Skripsi ini tidak akan terwujud tanpa adanya bantuan dan dukungan dari berbagai pihak. Ucapan terima kasih penulis sampaikan kepada :

1. Bapak Riduan Inayah dan Ibu Maimunah yang selalu memanjatkan doa untuk anak-anaknya serta dukungan semangat, nasehat dan motivasi yang luar biasa sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
2. Bapak Taghfirul Azhima Yoga Siswa, S. Kom, MM, M. Kom, yang telah bersedia meluangkan waktunya untuk membimbing kami, selalu memberikan penulis motivasi untuk dapat menyelesaikan skripsi dan kuliah tepat waktu, serta begitu banyak ilmu yang bapak bagikan terhadap penulis.
3. Ibu Asslia Johar Latipah, M.Cs., selaku ketua Program Studi S1 Teknik Informatika.
4. Bapak Prof. Ir. Sarjito, M.T.,Ph.D., selaku Dekan Fakultas Sains dan Teknologi
5. Bapak Prof. Dr. H. Bambang Setiaji, selaku Rektor Universitas Muhammadiyah Kalimantan Timur.
6. Teman - teman satu bimbingan penulis Muhammad Rizali Rahman, Nor Hikmah Sari, Anton Saputra, Muhammad Rhosyid Akhmad, Yoga Priantama, Asnur Karima, Agustya Nanda Pratiwi dan Rizky Aspiah yang selalu menemani bersama mengerjakan skripsi ini, serta memberikan dukungan dan motivasi demi meraih kesuksesan bersama.
7. Teman – Teman seperjuangan MAN 1, khususnya Dwiki Iswanto, Heri Setiawan, Muhammad Rizqy Ananda dan Muhammad Deni Faturrahman yang telah memberi banyak masukan dan dukungan dalam proses penyelesaian skripsi ini.

Semoga Allah SWT memberikan balasan atas segala jasa kebaikannya. Penulis menyadari skripsi ini tidak luput dari kesalahan maka dari itu kritik dan saran sangat diharapkan untuk menjadikan skripsi ini lebih baik lagi.

Samarinda, 30 Mei 2022

A handwritten signature in black ink, appearing to read 'Muhammad Norhalimi', written in a cursive style.

Muhammad Norhalimi

## ABSTRAK

Universitas Muhammadiyah Kalimantan Timur merupakan perguruan tinggi swasta yang pembiayaan kuliahnya dibebankan kepada mahasiswa melalui SPP, sehingga keterlambatan pembayaran kuliah tentu sangat berpengaruh terhadap operasional kampus. Maka perlu adanya pendekatan menggunakan data mining dengan metode klasifikasi algoritma *Naïve Bayes* untuk memprediksi keterlambatan pembayaran biaya kuliah di Universitas Muhammadiyah Kalimantan Timur. Tujuan dari penelitian ini menentukan indikator, mengimplementasikan algoritma *Naïve Bayes* dan mengevaluasi hasil kinerja algoritma *Naïve Bayes* dengan pembagian data 90:10 menggunakan *confusion matrix*. Data yang digunakan bersumber dari Bagian Administrasi Akademik dan Bagian Administrasi Keuangan yang berjumlah 12.408 data yang didalamnya memiliki beberapa atribut seperti penghasilan ayah, penghasilan ibu, pendidikan ayah, pendidikan ibu dan label *class* target tepat atau terlambat. Hasil dari implementasi algoritma *Naïve Bayes* menggunakan pembagian data 90:10 mendapatkan probabilitas tepat sebesar 2 dan probabilitas terlambat sebesar 4. Evaluasi menggunakan *confusion matrix* mendapatkan hasil *accuracy* sebesar 50,28%.

**Kata Kunci** : Prediksi, *Naïve Bayes*, *Confusion Matrix*

## **ABSTRACT**

*Muhammadiyah University of East Kalimantan is a private university whose tuition fees are charged to students through tuition fees, so the delay in tuition payments will certainly greatly affect campus operations. So there is a need for an approach using data mining with the Naïve Bayes algorithm classification method to predict late payment of tuition fees at the Muhammadiyah University of East Kalimantan. The purpose of this study is to determine indicators, implement the Naïve Bayes algorithm and evaluate the performance results of the Naïve Bayes algorithm with 90:10 data sharing using a confusion matrix. The data used is sourced from the Academic Administration Section and the Financial Administration Section, totaling 12,408 data in which it has several attributes such as father's income, mother's income, father's education, mother's education and class target right or late. The results of the implementation of the Naïve Bayes algorithm using 90:10 data division get an exact probability of 49,40% and a late probability of 50,60%. Evaluation using the confusion matrix got an accuracy of 50,28%.*

**Keywords :** Prediction, Naïve Bayes, Confusion Matrix



## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN JUDUL .....	i
HALAMAN PENGESAHAN .....	iii
PERNYATAAN KEASLIAN SKRIPSI .....	iv
PRAKATA.....	v
ABSTRAK.....	vii
<i>ABSTRACT</i> .....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR .....	xii
DAFTAR LAMPIRAN .....	xiii
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah .....	3
1.3. Tujuan Penelitian.....	3
1.4. Batasan Masalah .....	4
BAB 2 TINJAUAN PUSTAKA.....	5
2.1. Prediksi Keterlambatan Biaya Kuliah.....	5
2.2. <i>Data Mining</i> .....	6
2.3. Persiapan Data.....	7
2.4. Transformasi .....	8
2.5. Klasifikasi .....	8
2.6. <i>Algoritma Naive Bayes</i> .....	9
2.7. <i>Confusion Matrix</i> .....	12
BAB 3 METODE PENELITIAN .....	14
3.1. Tahapan Penelitian.....	14
3.2. Pengumpulan Data.....	14
3.3. Persiapan Data.....	15
3.4. Split Data .....	16

3.5. Permodelan.....	17
3.6. Evaluasi .....	17
BAB 4 HASIL DAN PEMBAHASAN .....	18
4.1 Data Penelitian.....	18
4.2 Persiapan Data.....	20
4.3 Pembagian Data 90:10.....	21
4.4 Permodelan Algoritma <i>Naïve Bayes</i> .....	21
4.6 Evaluasi menggunakan <i>confusion matrix</i> .....	26
BAB 5 PENUTUP .....	28
5.1. Kesimpulan.....	28
5.2. Saran.....	28
DAFTAR PUSTAKA.....	30
LAMPIRAN .....	34

## DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait Keterlambatan SPP .....	5
Tabel 2. 1 Penelitian Terkait Keterlambatan SPP (Lanjutan).....	6
Tabel 2. 2 Penelitian Terkait Algoritma <i>Naïve Bayes</i> .....	11
Tabel 2. 2 Penelitian Terkait Algoritma <i>Naïve Bayes</i> (Lanjutan).....	12
Tabel 2. 3 <i>Confusion Matrix</i> .....	12
Tabel 3. 1 Atribut Data Mahasiswa Dari Biro Administrasi Keuangan .....	14
Tabel 3. 1 Atribut Data Mahasiswa Dari Biro Administrasi Keuangan (Lanjutan) .....	15
Tabel 3. 2 Atribut Data Mahasiswa Dari Biro Administrasi Akademik .....	15
Tabel 3. 3 Atribut Setelah di seleksi.....	16
Tabel 3. 3 Atribut Setelah di seleksi (Lanjutan).....	17
Tabel 4. 1 Data Pembayaran SPP yang Bersumber dari BAK.....	19
Tabel 4. 2 Data Mahasiswa dari BAA Tahun 2017 - 2021 .....	20
Tabel 4. 3 Pembagian Data Training dan Testing.....	23
Tabel 4. 4 Data Training.....	23
Tabel 4. 5 Data Testing .....	23
Tabel 4. 6 Perhitungan Mean Tepat dan Terlambat.....	25
Tabel 4. 7 Standard Deviasi Tepat dan Terlambat .....	25
Tabel 4. 8 Perhitungan Nilai Gaussian Tepat .....	26
Tabel 4. 9 Perhitungan Nilai Gaussian Terlambat.....	27
Tabel 4. 10 Penentuan Nilai Akhir Tepat dan Terlambat.....	28
Tabel 4. 11 Nilai Aktual dan Prediksi .....	28
Tabel 4. 13 Evaluasi <i>Confusion Matrix</i> .....	29

## DAFTAR GAMBAR

Gambar 1. 1 Grafik Pembayaran SPP Periode 2019-2021 .....	2
Gambar 3. 1 Tahapan Penelitian .....	14
Gambar 4. 1 Pembagian data Training dan Testing.....	22

## **DAFTAR LAMPIRAN**

- Lampiran 1. Riwayat Hidup
- Lampiran 2. Data Bagian Biro Administrasi Keuangan
- Lampiran 3. Data Bagian Biro Administrasi Akademik
- Lampiran 4. Integrasi Data
- Lampiran 5. Pembersihan data
- Lampiran 6. Reduksi data
- Lampiran 7. Seleksi data
- Lampiran 8. Transformasi Data
- Lampiran 9. Coding
- Lampiran 10. Surat Ijin Penelitian Biro Administrasi Keuangan
- Lampiran 11. Surat Ijin Penelitian Biro Administrasi Akademik
- Lampiran 12. Surat Keterangan Tidak Uji Validitas
- Lampiran 13. Lembar Bimbingan Skripsi
- Lampiran 14. Hasil Uji Plagiasi