

**KLASIFIKASI KELULUSAN TEPAT WAKTU MAHASISWA UIN SUNAN  
AMPEL SURABAYA DENGAN MENGGUNAKAN METODE *NAÏVE BAYES***

**SKRIPSI**



**UIN SUNAN AMPEL  
S U R A B A Y A**

**Disusun Oleh:**

**JAMIATUS SHOLIAH**

**NIM: H72215015**

**PROGRAM STUDI MATEMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS ISLAM NEGERI SUNAN AMPEL**

**SURABAYA**

**2020**

## PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : Jamiatus Sholihah

NIM : H72215015

Program Studi : Matematika

Angkatan : 2015

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi yang berjudul “KLASIFIKASI KELULUSAN TEPAT WAKTU MAHASISWA UIN SUNAN AMPEL DENGAN MENGGUNAKAN METODE *NAÏVE BAYES* ”. Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 27 Juli 2020

Yang menyatakan,



(Jamiatus Sholihah)

NIM: H72215015

## LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh :

NAMA : Jamiatus Sholihah

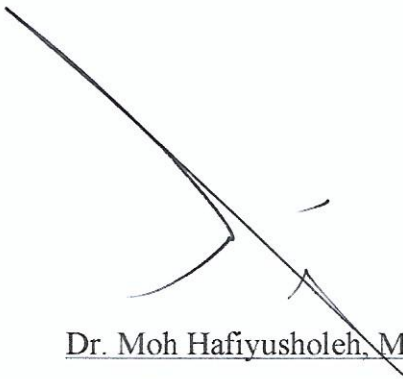
NIM : H72215015

JUDUL : Klasifikasi Kelulusan Tepat Waktu Mahasiswa UIN Sunan  
Ampel Surabaya dengan Menggunakan Metode *Naïve Bayes*.

Telah diperiksa dan disetujui untuk diujikan.

Surabaya, 6 Agustus 2020

Dosen Pembimbing 1



Dr. Moh Hafiyusholeh, M.Si, M.PMat

NIP: 198002042014031001

Dosen Pembimbing 2



Dian Candra Rini N, M. Kom

NIP: 198511242014032001

## PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh :

NAMA : Jamiatus Sholihah

NIM : H72215015

JUDUL : Klasifikasi Kelulusan Tepat Waktu Mahasiswa UIN Sunan Ampel Surabaya dengan Menggunakan Metode *Naïve Bayes*.


Telah dipertahankan di depan tim penguji skripsi

Pada hari Kamis, Tanggal 10 Agustus 2020


Mengesahkan,

Tim Penguji


Penguji I

  
Dr. Moh Hafiyusholeh, M.Si, M.PMat  
NIP. 198002042014031001


Penguji II

  
Dian Candra Rini N. M. Kom  
NIP. 198511242014032001

Penguji III

  
Nurissaidah Ulinnuha, M.Kom  
NIP. 199011022014032004

Penguji IV

  
Wika Dianita Utami, M.Sc  
NIP. 199206102018012003

Mengetahui,

Plt. Fakultas Sains dan Teknologi

UIN Sunan Ampel Surabaya



Plt. Dekan Karmatur Rusydiyah, M.Ag

NIP. 197312272005012003



UIN SUNAN AMPEL  
S U R A B A Y A

KEMENTERIAN AGAMA  
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA  
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300  
E-Mail: perpustakaan@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI  
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Jamiatus Sholihah  
NIM : H72215015  
Fakultas/Jurusan : SAINTEK/MATEMATIKA  
E-mail address : Jamiatussholihah33@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

Skripsi  Tesis  Desertasi  Lain-lain (.....)  
yang berjudul :

KLASIFIKASI KELULUSAN TEPAT WAKTU MAHASISWA UIN SUNAN AMPEL  
SURABAYA DENGAN MENGGUNAKAN NAIVE BAYES

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 30 Januari 2023

Penulis

  
( Jamiatus Sholihah )  
nama terang dan tanda tangan



## ABSTRAK

### **KLASIFIKASI KELULUSAN TEPAT WAKTU MAHASISWA UIN SUNAN AMPEL SURABAYA DENGAN MENGGUNAKAN METODE NAÏVE BAYES**

Data kelulusan mahasiswa sangat penting untuk universitas maupun program studi karena data tersebut merupakan salah satu instrumen dalam penilaian proses akreditasi universitas maupun program studi. Data kelulusan di uinsa sendiri pada tahun 2012 mempunyai persentase 12% hal ini menunjukkan bahwa antara quota masuk pada tahun ajaran baru dan tingkat kelulusan sangatlah tidak seimbang. Oleh karena itu dibutuhkannya sebuah sistem yang dapat mengklasifikasikan kelulusan mahasiswa sehingga kita dapat mengetahui pola data seperti apa yang menyebabkan ketidakseimbangan antara quota masuk dan data kelulusan. Pada penelitian ini akan menggunakan metode Naïve Bayes untuk membuat sistem klasifikasi antara mahasiswa lulus tepat waktu dan mahasiswa lulus tidak tepat waktu. Hasil dari penelitian ini ialah dengan menggunakan perbandingan data *training* 80% dan data *testing* 20% mempunyai akurasi tertinggi yaitu 80,43% dengan 58 mahasiswa yang tepat terklasifikasikan sebagai mahasiswa lulus tepat waktu dan 920 mahasiswa yang diprediksikan sebagai mahasiswa yang lulus tidak tepat waktu.

Kata Kunci: *kelulusan mahasiswa, klasifikasi, Naïve Bayes*

UIN SUNAN AMPEL  
S U R A B A Y A

## **ABSTRACT**

### ***CLASSIFICATION OF TIME STUDENT GRADUATION OF UIN SUNAN AMPEL SURABAYA USING NAÏVE BAYES METHOD***

*Student graduation data is very important for the university and the study program because the data are an instrument in assessing accreditation process of the university and the study program. In addition, student graduation data can be used to evaluate and repair the learning system so that it can improve the quality of the university and the study program. To find out more information about the data, it is necessary to process the data and the analysis results from those processing. As in this study, which discusses the processing of student graduation data using the NEIVA BAYE method to create a classification system between students graduating on time and students graduating not on time. By using different data comparisons, it can be seen which data comparison has the maximum accuracy. The result is the comparison of data with training data 80% and testing data 20% has the highest accuracy, i.e 80.43%.*

*Keyword: Student graduation, classification, Naïve Bayes*



UIN SUNAN AMPEL  
S U R A B A Y A

## DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING .....	i
PENGESAHAN TIM PENGUJI SKRIPSI.....	ii
PERNYATAAN KEASLIAN.....	iii
MOTO.....	iv
HALAMAN PERSEMBAHAN .....	v
KATA PENGANTAR .....	vi
ABSTRAK.....	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan Penelitian .....	5
1.4 Manfaat Penelitian .....	5
BAB II LANDASAN TEORI.....	7
2.1 Kelulusan Mahasiswa .....	7
2.2 Klasifikasi .....	7
2.3 Algoritma <i>Naïve Bayes</i> Classifier .....	8
2.4 Pengujian Model Klasifikasi.....	12
Tabel 2.1 <i>Confussion Matrix</i> .....	12
BAB III METODE PENELITIAN.....	15
3.1 Jenis Penelitian.....	15
3.2 Sumber Data.....	15
3.3 Pengolahan Data .....	15



3.4 Tahapan Penelitian.....	16
Gambar 3.1 Flowchart rangkaian penelitian .....	16
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>18</b>
4.1 Perhitungan Metode <i>Naïve Bayes</i> .....	18
4.1.1 Data <i>Training</i> 90% dan Data <i>Testing</i> 10% .....	18
Tabel 4.1 nilai rata-rata dan standar deviasi kelas 1 .....	19
Tabel 4.2 nilai rata-rata dan standar deviasi kelas 0 .....	19
Tabel 4.3 nilai distribusi gaussian .....	20
4.1.2 Data <i>Training</i> 85% dan Data <i>Testing</i> 15% .....	21
Tabel 4.4 nilai rata-rata dan standar deviasi kelas 1 .....	22
Tabel 4.5 nilai rata-rata dan standar deviasi kelas 0 .....	22
Tabel 4.6 nilai distribusi gaussian .....	23
4.1.3 Data <i>Training</i> 80% dan Data <i>Testing</i> 20% .....	24
Tabel 4.7 nilai rata-rata dan standar deviasi kelas 1 .....	24
Tabel 4.8 nilai rata-rata dan standar deviasi kelas 0 .....	24
Tabel 4.9 nilai distribusi gaussian .....	25
4.1.4 Data <i>Training</i> 75% dan Data <i>Testing</i> 25% .....	26
Tabel 4.10 nilai rata-rata dan standar deviasi kelas 1 .....	27
Tabel 4.11 nilai rata-rata dan standar deviasi kelas 0 .....	27
Tabel 4.12 nilai distribusi gaussian .....	27
4.1.5 Data <i>Training</i> 70% dan Data <i>Testing</i> 30% .....	28
Tabel 4.13 nilai rata-rata dan standar deviasi kelas 1 .....	29
Tabel 4.14 nilai rata-rata dan standar deviasi kelas 0 .....	29
Tabel 4.15 nilai distribusi gaussian .....	30
4.1.6 Data <i>Training</i> 65% dan Data <i>Testing</i> 35% .....	31
Tabel 4.16 nilai rata-rata dan standar deviasi kelas 1 .....	31
Tabel 4.17 nilai rata-rata dan standar deviasi kelas 0 .....	31

Tabel 4.18 nilai distribusi gaussian .....	32
4.2 Hasil Evaluasi dan Akurasi .....	33
4.2.1 <i>Confussion matrix</i> .....	34
4.2.1.1 Pengujian 1.....	34
Tabel 4.19 <i>Confussion Matrix</i> pengujian 1 .....	34
4.2.1.2 Pengujian 2.....	35
Tabel 4.20 <i>Confussion Matrix</i> pengujian 2 .....	35
4.2.1.3 Pengujian 3.....	36
Tabel 4.21 <i>Confussion Matrix</i> pengujian 3 .....	36
4.2.1.4 Pengujian 4.....	37
Tabel 4.22 <i>Confussion Matrix</i> pengujian 4 .....	38
4.2.1.5 Pengujian 5.....	39
Tabel 4.23 <i>Confussion Matrix</i> pengujian 5 .....	39
4.2.1.6 Pengujian 6.....	40
Tabel 4.24 <i>Confussion Matrix</i> pengujian 6 .....	40
Tabel 4.24 Perbandingan pengujian hasil klasifikasi .....	41
4.3 Analisa Hasil.....	41
BAB V PENUTUP.....	43
5.1 Kesimpulan .....	43
5.2 Saran .....	44
DAFTAR PUSTAKA .....	70

## DAFTAR TABEL

Tabel 2.1 <i>Confussion Matrix</i> .....	12
Tabel 4.1 nilai rata-rata dan standar deviasi kelas 1 .....	19
Tabel 4.2 nilai rata-rata dan standar deviasi kelas 0 .....	19
Tabel 4.3 nilai distribusi gaussian .....	20
Tabel 4.4 nilai rata-rata dan standar deviasi kelas 1 .....	22
Tabel 4.5 nilai rata-rata dan standar deviasi kelas 0 .....	22
Tabel 4.6 nilai distribusi gaussian .....	23
Tabel 4.7 nilai rata-rata dan standar deviasi kelas 1 .....	24
Tabel 4.8 nilai rata-rata dan standar deviasi kelas 0 .....	24
Tabel 4.9 nilai distribusi gaussian .....	25
Tabel 4.10 nilai rata-rata dan standar deviasi kelas 1 .....	27
Tabel 4.11 nilai rata-rata dan standar deviasi kelas 0 .....	27
Tabel 4.12 nilai distribusi gaussian .....	27
Tabel 4.13 nilai rata-rata dan standar deviasi kelas 1 .....	29
Tabel 4.14 nilai rata-rata dan standar deviasi kelas 0 .....	29
Tabel 4.15 nilai distribusi gaussian .....	30
Tabel 4.16 nilai rata-rata dan standar deviasi kelas 1 .....	31
Tabel 4.18 nilai distribusi gaussian .....	32
Tabel 4.19 <i>Confussion Matrix</i> pengujian 1 .....	34
Tabel 4.20 <i>Confussion Matrix</i> pengujian 2 .....	35
Tabel 4.21 <i>Confussion Matrix</i> pengujian 3 .....	36
Tabel 4.22 <i>Confussion Matrix</i> pengujian 4 .....	38
Tabel 4.23 <i>Confussion Matrix</i> pengujian 5 .....	39

Tabel 4.24 *Confussion Matrix* pengujian 6..... 40

Tabel 4.24 Perbandingan pengujian hasil klasifikasi..... 41



UIN SUNAN AMPEL  
S U R A B A Y A

## DAFTAR GAMBAR

Gambar 3.1 *Flowchart* Rangkaian Penelitian ..... 12



## DAFTAR PUSTAKA

- BAN-PT. (2008). *Akreditasi Program Studi Sarjana Program Studi Peternakan* (Issue april). BAN-PT.
- Dewi, S. (2016). Komparasi 5 Metode Algoritma Klasifikasi Data Mining pada Prediksi Keberhasilan Pemasaran Produk Layanan Perbankan. *Jurnal Techno Nusa Mandiri*, XIII(1), 60–66.
- Faisal, M. R. (n.d.). Seri Belajar Data Science : Klasifikasi dengan Bahasa Pemrograman R. In *Indonesia Net Developer Comunity*.
- Farida, Y., & Ulinuha, N. (2018). *Kalsifikasi Mahasiswa Penerima Program Beasiswa Bidik Misi Menggunakan Niave Bayes*. 04(01), 17–22.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (Third Edit).
- Hanggara, A. H., Awang, I. S., & Bejo. (2016). Pengaruh Model Pembelajaran Numbered Heads Together Terhadap Hasil Belajar. *Jurnal Pendidikan Dasar Perkasa*, 2(November).
- Irfan, M. (2015). *Analisa Pola Asosiasi Jalur Masuk terhadap Kelulusan Mahasiswa dengan Menggunakan Metode FOLD-GROWTH*. IX(2).
- Listiowarni, I., & Ramadhani, N. (2019). Implementasi Naïve Bayesian dengan Laplacian Smoothing untuk Peminatan dan Lintas Minat Siswa SMAN 5 Pamekasan. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 8(2), 124. <https://doi.org/10.32736/sisfokom.v8i2.652>
- Nugroho, M. F., & Wibowo, S. (2017). Fitur Seleksi Forward Selection untuk Menentukan Atribut yang Berpengaruh pada Klasifikasi Kelulusan Mahasiswa Fakultas Ilmu Komputer UNAKI Semarang Menggunakan Algoritma Naive Bayes. *Jurnal Informatika Upgris*, 3(1), 63–70. <https://doi.org/10.26877/jiu.v3i1.1669>
- Prasetyo, E. (2012). *Data Mining : Konsep dan Aplikasi Menggunakan MATLAB* (N. WK (ed.); I). ANDI Yogyakarta.
- Pujianto, U., Widiiyaningtyas, T., Prestya, D. D., & Romadhon, B. (2017). Penerapan Algoritma Naïve Bayes Classifier untuk Klasifikasi Judul Skripsi dan Tugas Akhir Berdasarkan Kelompok Bidang Keahlian. *Jurnal Teknologi*



*Elektro Dan Kejuruan*, 27(1), 79–92.

Putri, R. P. S., & Waspada, I. (2018). Penerapan Algoritma C4.5 pada Aplikasi Prediksi Kelulusan Mahasiswa Prodi Informatika. *Khazanah Informatika: Jurnal Ilmu Komputer Dan Informatika*, 4(1), 1.

<https://doi.org/10.23917/khif.v4i1.5975>

Santoso, H. (2017). Pendidikan Karakter untuk Menyiapkan Generasi Indonesia Berkemajuan. *Seminar Nasional Pendidikan*, 404–409.

UINSA. (2016). *Pedoman Akademik UIN Sunan Ampel Surabaya 2016*.

Vercellis, C. (2009). Business Intelligence: Data Mining and Optimization for Decision Making. In *A John Wiley and Sons, Ltd.* John Wiley and Sons Ltd.

<https://doi.org/10.1002/9780470753866>



UIN SUNAN AMPEL  
S U R A B A Y A