

**IMPLEMENTASI *EDUCATIONAL DATA MINING*
MENGUNAKAN NEURAL NETWORK UNTUK
MEMPREDIKSI WAKTU KELULUSAN MAHASISWA (STUDI
KASUS: UNIVERSITAS MULTIMEDIA NUSANTARA)**



SKRIPSI

**BILL KIKI
00000043854**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2024**

**IMPLEMENTASI *EDUCATIONAL DATA MINING*
MENGUNAKAN NEURAL NETWORK UNTUK
MEMPREDIKSI WAKTU KELULUSAN MAHASISWA (STUDI
KASUS: UNIVERSITAS MULTIMEDIA NUSANTARA)**



SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer (S.Kom.)

Bill Kiki

00000043854

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

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Bill Kiki
Nomor Induk Mahasiswa : 00000043854
Program studi : Sistem Informasi

Skripsi dengan judul:

“Implementasi *Educational Data Mining* Menggunakan Neural Network untuk Memprediksi Waktu Kelulusan Mahasiswa (Studi Kasus: Universitas Multimedia Nusantara)”

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, 30 Mei 2024


Bill Kiki

UNIVERSITAS
MULTIMEDIA
NUSANTARA

MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

“Implementasi Educational Data Mining Menggunakan Neural Network untuk
Memprediksi Waktu Kelulusan Mahasiswa (Studi Kasus: Universitas Multimedia
Nusantara)”

Oleh


Nama : Bill Kiki
NIM : 0000043854
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Kamis, 30 Mei 2024
Pukul 13.00 s.d 15.00 dan dinyatakan
LULUS

Dengan susunan penguji sebagai berikut.


Ketua Sidang

Penguji


Ririn Ikana Desanti, S.Kom, M.Kom.
0313058001


Dinar Ajeng Kristiyanti, S.Kom., M.Kom.
0330128801

Pembimbing


Dr. Friska Natalia, S.Kom., M.T
0306128307

Ketua Program Studi Sistem Informasi


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

iii

Implementasi Educational Data..., Bill Kiki, Universitas Multimedia Nusantara

MULTIMEDIA
NUSANTARA

iii

Implementasi Educational Data..., Bill Kiki, Universitas Multimedia Nusantara

LEMBAR PERSETUJUAN PUBLIKASI

KARYA ILMIAH MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : Bill Kiki
NIM : 00000043854
Program Studi : Sistem Informasi
Jenjang : S1
Judul Karya Ilmiah : Implementasi Educational Data Mining Menggunakan Neural Network Untuk Memprediksi Waktu Kelulusan Mahasiswa (Studi Kasus: Universitas Multimedia Nusantara)

Menyatakan dengan sesungguhnya bahwa saya bersedia:

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial. Saya tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: Dalam proses pengajuan penerbitan ke dalam jurnal/konferensi nasional/internasional (dibuktikan dengan letter of acceptance)*.

Tangerang, 30 Mei 2024

UNIVERSITAS
MULTIMEDIA
NUSANTARA


Bill Kiki

* Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

KATA PENGANTAR

Puja-Puji Syukur saya ucapkan atas terselesaikannya Proposal Skripsi ini yang berjudul “Implementasi *Educational Data Mining* Menggunakan Neural Network untuk Memprediksi Waktu Kelulusan Mahasiswa (Studi Kasus: Universitas Multimedia Nusantara)” yang dilakukan untuk memenuhi salah satu dari syarat kelulusan Sarjana (S1) dan mendapatkan gelar S.Kom Jurusan Sistem Informasi Universitas Multimedia Nusantara. Tanpa bantuan, arahan, serta bimbingan dari berbagai pihak, penyelesaian penulisan proposal skripsi ini akan menjadi lebih rumit dan sulit. Saya ingin mengucapkan terima kasih pada:

1. Dr. Ninok Leksono, M.A., selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, ST, M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Dr. Friska Natalia, S.Kom., M.T., sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tesis ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.

Semoga karya ilmiah ini dapat berdampak positif bagi para pembaca dan menjadi sumber informasi serta wawasan yang bermanfaat.

Tangerang, 13 Mei 2024

UNIVERSITAS
MULTIMEDIA
NUSANTARA



Bill Kiki

IMPLEMENTASI *EDUCATIONAL DATA MINING*
MENGGUNAKAN NEURAL NETWORK UNTUK
MEMPREDIKSI WAKTU KELULUSAN MAHASISWA (STUDI
KASUS: UNIVERSITAS MULTIMEDIA NUSANTARA)

Bill Kiki

ABSTRAK

Dalam beberapa tahun terakhir, pertumbuhan jumlah lembaga pendidikan tinggi terus meningkat. Hal ini tidak terlepas dari persaingan antar perguruan tinggi. Salah satu ukuran kualitas pendidikan tinggi adalah tingkat kelulusan mahasiswa, yang mencakup masa studi sebagai salah satu komponennya. Bagi lembaga pendidikan, mahasiswa yang lulus tidak tepat waktu tentu akan mempengaruhi citra dan akreditasi lembaga pendidikan tersebut. Maka dari itu, penelitian ini dilakukan dengan tujuan untuk membuat model prediksi waktu kelulusan mahasiswa alternatif solusi untuk meningkatkan kesadaran mahasiswa serta sebagai salah satu acuan bagi dosen pembimbing menyesuaikan pemberian bimbingan akademik kepada para mahasiswa bimbingan. Penelitian ini menggunakan tiga algoritma klasifikasi yaitu Naïve Bayes, Support Vector Machine dan Neural Network yang dioptimasi menggunakan algoritma Particle Swarm Optimization. Pengoptimalan dilakukan karena parameter default dari ketiga algoritma klasifikasi cenderung rentan terhadap serangan adversarial. *Framework* CRISP-DM digunakan sebagai alur penelitian dengan 6 tahapan, yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan data, evaluasi, dan penerapan. Metrik evaluasi yang digunakan adalah akurasi, presisi, sensitivitas, spesifisitas, dan *f1-score* untuk menilai algoritma terbaik. Hasil penelitian menunjukkan bahwa performa model terbaik diraih oleh algoritma Neural Network setelah optimasi yang mendapat peningkatan akurasi sebesar 8%, yaitu dari 78% menjadi 86%. Neural Network menjadi algoritma terbaik dengan nilai akurasi, presisi, sensitivitas, spesifisitas, dan *f1-score* secara berturut-turut sebesar 0.86, 0.86, 0.86, 0.98, dan 0.86. Hasil model akan diimplementasikan untuk melakukan prediksi terhadap data mahasiswa berstatus aktif dan ditampilkan dalam bentuk *website*.

Kata kunci: Naïve Bayes, Neural Network, Particle Swarm Optimization, Prediksi Waktu Kelulusan Mahasiswa, Support Vector Machine

**IMPLEMENTATION OF EDUCATIONAL DATA MINING
USING NEURAL NETWORK FOR PREDICTING STUDENT
GRADUATION TIME (CASE STUDY: MULTIMEDIA
NUSANTARA UNIVERSITY)**

Bill Kiki

ABSTRACT (English)

In recent years, the growth in the number of higher education institutions has continued to increase. This is not separate from the competition among universities. One measure of the quality of higher education is the graduation rate of students, which includes the duration of study as one of its components. For educational institutions, students who do not graduate on time can certainly affect the institution's image and accreditation. Therefore, this research aims to create a predictive model for student graduation times as an alternative solution to raise student awareness and as a reference for academic advisors to tailor their guidance to their advisees. This research employs three classification algorithms: Naïve Bayes, Support Vector Machine, and Neural Network, optimized using the Particle Swarm Optimization algorithm. Optimization is carried out because the default parameters of the three classification algorithms tend to be vulnerable to adversarial attacks. The CRISP-DM framework is used as the research workflow with six stages: business understanding, data understanding, data preparation, data modeling, evaluation, and deployment. The evaluation metrics used are accuracy, precision, sensitivity, specificity, and f1-score to assess the best algorithm. The research results show that the best model performance is achieved by the Neural Network algorithm after optimization, which sees an accuracy increase of 8%, from 78% to 86%. Neural Network becomes the best algorithm with accuracy, precision, sensitivity, specificity, and f1-score values of 0.86, 0.86, 0.86, 0.98, and 0.86, respectively. The model results will be implemented to make predictions for active student data and displayed in the form of a website.

Keywords: *Naïve Bayes, Neural Network, Particle Swarm Optimization, Student Graduation Time Prediction, Support Vector Machine*

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT Error! Bookmark not defined.	
HALAMAN PENGESAHANError! Bookmark not defined.	
LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA	iv
KATA PENGANTAR	v
ABSTRAK	vi
ABSTRACT (English)	vii
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat Penelitian	5
1.4.1 Tujuan Penelitian	5
1.4.2 Manfaat Penelitian	5
1.5 Sistematika Penulisan	6
BAB II LANDASAN TEORI	8
2.1 Penelitian Terdahulu	8
2.2 Tinjauan Teori	13
2.2.1 Perguruan tinggi	13
2.2.2 Kelulusan	13
2.2.3 Data Mining	14
2.2.4 Educational Data Mining	14
2.3 Framework, Algoritma, Persiapan Data dan Metode Evaluasi	16
2.3.1 Framework	16
2.3.1.1 Framework CRISP-DM	16
2.3.1.2 Framework KDD	19
2.3.1.3 Framework SEMMA	19

2.3.2	Algoritma	19
2.3.2.1	Neural-Network (NN)	19
2.3.2.2	Naïve Bayes (NB).....	20
2.3.2.3	Support Vector Machine (SVM).....	21
2.3.2.4	Particle Swarm Optimization (PSO)	22
2.3.3	Persiapan Data	24
2.3.3.1	<i>Data Imputation</i>	24
2.3.3.2	<i>K-fold Cross Validation</i>	24
2.3.4	Evaluasi.....	24
2.4	Alat Penelitian	28
2.4.1	Python	28
2.4.2	HTML	28
2.4.3	CSS	28
2.4.4	PHP	29
2.4.5	Visual Studio Code.....	29
BAB III METODOLOGI PENELITIAN		31
3.1	Gambaran Umum Objek Penelitian.....	31
3.2	Metode Penelitian.....	31
3.2.1	Alur Penelitian	31
3.2.2	Metode <i>Data Mining</i>	33
3.2.2.1	<i>Business Understanding</i>	34
3.2.2.2	<i>Data Understanding</i>	35
3.2.2.3	<i>Data Preparation</i>	36
3.2.2.4	<i>Modeling</i>	37
3.2.2.5	<i>Evaluation</i>	38
3.2.2.6	<i>Deployment</i>	39
3.3	Teknik Pengumpulan Data	41
3.3.1	Populasi dan Sampel.....	42
3.3.2	Periode Pengambilan Data	42
3.4	Variabel Penelitian.....	42
3.5	Teknik Analisis Data.....	43
BAB IV ANALISIS DAN HASIL PENELITIAN		45
4.1	<i>Business Understanding</i>	45
4.2	<i>Data Understanding</i>	46

4.2.1 Deskripsi Data	47
4.2.2 <i>Exploratory Data Analysis (EDA)</i>	48
4.3 <i>Data Preparation</i>	55
4.3.1 Penghapusan Kolom dan Anomali Data.....	55
4.3.2 <i>Re-format Data</i>	56
4.3.3 Augmentasi Data.....	58
4.3.4 Pemisahan Data.....	59
4.3.5 Normalisasi Data dan <i>One-Hot Encoding</i> pada Label.....	60
4.4 <i>Data Modeling</i>	61
4.4.1 Algoritma Naïve Bayes	61
4.4.2 Algoritma Support Vector Machine	63
4.4.3 Algoritma Neural Network	65
4.5 <i>Evaluation</i>	67
4.6 <i>Deployment</i>	74
4.6.1 Penyimpanan Model dan Prediksi Data Baru	75
4.6.2 Penyimpanan Data pada <i>Database</i>	76
4.6.3 Pembuatan <i>Website</i>	77
4.7 Hasil dan Diskusi	80
BAB V SIMPULAN DAN SARAN	86
5.1 Simpulan.....	86
5.2 Saran.....	87
DAFTAR PUSTAKA	88
LAMPIRAN	94

UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	8
Tabel 2.2 Pseudocode Algoritma Neural Network (NN).....	20
Tabel 2.3 Pseudocode Algoritma Naïve Bayes (NB).....	21
Tabel 2.4 Pseudocode Algoritma Support Vector Machine (SVM)	22
Tabel 2.5 Pseudocode Algoritma Particle Swarm Optimization (PSO)	23
Tabel 3.1 Perbandingan CRISP-DM, SEMMA, dan KDD.....	34
Tabel 3.2 Perbandingan Algoritma Klasifikasi	39
Tabel 3.3 Perbandingan R dan Python [67],[68],[69].....	44
Tabel 4.1 Deskripsi Kolom Data Mahasiswa UMN	48
Tabel 4.2 Hasil Pemisahan Data	60
Tabel 4.3 Perbandingan Hasil Performa Setiap Algoritma Sebelum Optimasi	69
Tabel 4.4 Perbandingan Hasil Performa Setiap Algoritma Setelah Optimasi	69
Tabel 4.5 Hasil Performa Algoritma Terbaik Setiap Lipatan	74
Tabel 4.6 Perbandingan Hasil Penelitian dengan Penelitian Terdahulu	81
Tabel 4.7 Pernyataan dan Nilai Jawaban Responden Mahasiswa.....	83
Tabel 4.8 Pernyataan dan Nilai Jawaban Responden Dosen Pembimbing atau Staf Akademik	83



DAFTAR GAMBAR

Gambar 2.1 Framework CRISP-DM.....	17
Gambar 3.1 Alur Penelitian.....	32
Gambar 4.1 Struktur Data Mahasiswa Lulus dan Aktif.....	47
Gambar 4.2 Kode Struktur Data.....	47
Gambar 4.3 Tampilan Sampel Data Mahasiswa Lulus dan Aktif.....	47
Gambar 4.4 Kode Statistika Deskriptif.....	48
Gambar 4.5 Statistika Deskriptif Variabel Numerik.....	49
Gambar 4.6 Kode Fungsi Visualisasi Variabel Kategorik.....	50
Gambar 4.7 Visualisasi Data Mahasiswa Berdasarkan Angkatan.....	50
Gambar 4.8 Visualisasi Data Mahasiswa Berdasarkan Tahun Lulus.....	51
Gambar 4.9 Visualisasi Data Mahasiswa Berdasarkan Masa Studi Semester.....	52
Gambar 4.10 Visualisasi Data Mahasiswa Berdasarkan Semester.....	53
Gambar 4.11 Visualisasi Data Mahasiswa Berdasarkan Semester.1.....	54
Gambar 4.12 Visualisasi Data Mahasiswa Berdasarkan Program Studi.....	54
Gambar 4.13 Visualisasi Data Mahasiswa Berdasarkan Status.....	55
Gambar 4.14 Penghapusan Kolom dan Penanganan Anomali.....	56
Gambar 4.15 Format Awal Data.....	57
Gambar 4.16 Transformasi Format Data.....	57
Gambar 4.17 Transformasi Urutan Kolom Data.....	57
Gambar 4.18 Format Setelah Transformasi Data.....	58
Gambar 4.19 Penambahan Kolom 1.....	58
Gambar 4.20 Penambahan Kolom 2.....	58
Gambar 4.21 Augmentasi Data 2.....	59
Gambar 4.22 Augmentasi Data 3.....	59
Gambar 4.23 Normalisasi Data.....	60
Gambar 4.24 One-Hot Encoding pada Label.....	61
Gambar 4.25 Import Library Algoritma Naïve Bayes.....	61
Gambar 4.26 Kode Model Algoritma Naïve Bayes Sebelum Optimasi.....	61
Gambar 4.27 Akurasi Algoritma Naïve Bayes Sebelum Optimasi.....	62
Gambar 4.28 Import Library Algoritma PSO.....	62
Gambar 4.29 Optimasi Model Algoritma Naïve Bayes dengan PSO.....	62
Gambar 4.30 Parameter Terbaik dan Akurasi Algoritma Naïve Bayes Setelah Optimasi.....	62
Gambar 4.31 Kode Model Algoritma Naïve Bayes Setelah Optimasi.....	63
Gambar 4.32 Import Library Algoritma SVM.....	63
Gambar 4.33 Kode Model Algoritma SVM Sebelum Optimasi.....	63
Gambar 4.34 Akurasi Algoritma SVM Sebelum Optimasi.....	63
Gambar 4.35 Optimasi Model Algoritma SVM dengan PSO.....	64
Gambar 4.36 Parameter Terbaik dan Akurasi Algoritma SVM Setelah Optimasi.....	64
Gambar 4.37 Kode Model Algoritma SVM Setelah Optimasi.....	65
Gambar 4.38 Import Library Algoritma NN.....	65
Gambar 4.39 Kode Model Algoritma NN Sebelum Optimasi.....	66
Gambar 4.40 Akurasi Algoritma NN Sebelum Optimasi.....	66
Gambar 4.41 Optimasi Model Algoritma NN dengan PSO.....	67

Gambar 4.42 Parameter Terbaik dan Akurasi Algoritma NN Setelah Optimasi ..	67
Gambar 4.43 Kode Model Algoritma NN Setelah Optimasi	67
Gambar 4.44 Import Library Fungsi Evaluasi	68
Gambar 4.45 Perhitungan Fungsi Evaluasi Spesifisitas.....	68
Gambar 4.46 Confusion Matrix Agloritma NB	71
Gambar 4.47 Confusion Matrix Algoritma SVM	72
Gambar 4.48 Confusion Matrix Algoritma NN	73
Gambar 4.49 Statistika Deskriptif Performa Setiap Lipatan Algoritma Terbaik..	74
Gambar 4.50 Penyimpanan Model.....	75
Gambar 4.51 Prediksi Data Baru	76
Gambar 4.52 Penambahan Kolom Dosen Pembimbing.....	76
Gambar 4.53 Insialisasi Koneksi Database	76
Gambar 4.54 Pembuatan Tabel dan Inseri Data	77
Gambar 4.55 Database Data Mahasiswa.....	77
Gambar 4.56 Tampilan Login Website	78
Gambar 4.57 Tampilan Halaman Website Mahasiswa	79
Gambar 4.58 Login Website Dosen Pembimbing	79
Gambar 4.59 Tampilan Halaman Pertama Website Dosen Pembimbing	80
Gambar 4.60 Tampilan Halaman Kedua Website Dosen Pembimbing.....	80

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR RUMUS

Rumus 2.1 Rumus Algoritma Neural Network [33]	20
Rumus 2.2 Rumus Algoritma Naïve Bayes [35].....	21
Rumus 2.3 Rumus Algoritma Support Vector Machine [36]	22
Rumus 2.4 Rumus Algoritma Particle Swarm Optimization [36]	23
Rumus 2.5 Rumus Metrik Akurasi [43]	25
Rumus 2.6 Rumus Metrik Presisi [43]	25
Rumus 2.7 Rumus Metrik Spesifisitas [43]	26
Rumus 2.8 Rumus Metrik Sensitivitas [43]	26
Rumus 2.9 Rumus Metrik f1-score [44]	27



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran A Pengecekan Turnitin	94
Lampiran B Formulir Konsultasi Skripsi	111
Lampiran C Surat Perizinan Pengambilan Data	112
Lampiran D Surat Pengambilan Data	114
Lampiran E Kuesioner <i>User Acceptance Testing</i>	115
Lampiran F Kelompok Data Bersih	118
Lampiran G Pertemuan Bersama Departemen BIA dan IT UMN	119



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA