



**PENERAPAN KLASIFIKASI PROSES SELEKSI PENERIMAAN MAHASISWA
MAGANG KAMPUS MERDEKA DI PT. XYZ**

SKRIPSI

**MUHAMMAD RAYHAN ATHAURRAHMAN
NIM. 2110512160**

**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" Jakarta
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI S1 SISTEM INFORMASI
2025**

**PENERAPAN KLASIFIKASI PROSES SELEKSI PENERIMAAN MAHASISWA
MAGANG KAMPUS MERDEKA DI PT. XYZ**



**MUHAMMAD RAYHAN ATHAURRAHMAN
NIM. 2110512160**

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
JAKARTA
2025**

PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini:

Nama : Muhammad Rayhan Athaurrahman
NIM : 2110512160
Program Studi : S1 Sistem Informasi
Fakultas : Fakultas Ilmu Komputer
Universitas : Universitas Pembangunan Nasional "Veteran" Jakarta

Dengan ini menyatakan bahwa karya yang saya buat berjudul "PENERAPAN KLASIFIKASI PROSES SELEKSI PENERIMAAN MAHASISWA MAGANG KAMPUS MERDEKA DI PT. XYZ" adalah merupakan hasil kerja asli saya sendiri dan belum pernah dipublikasikan atau diserahkan kepada pihak lain untuk tujuan yang sama. Saya juga menyatakan bahwa seluruh referensi yang digunakan dalam karya ini telah disebutkan dengan jelas dan sesuai dengan ketentuan yang berlaku.

Jika di kemudian hari terbukti bahwa karya ini mengandung unsur plagiasi atau melanggar hak cipta, saya bersedia menerima sanksi sesuai dengan peraturan yang berlaku di Universitas Pembangunan Nasional "Veteran" Jakarta.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya tanpa ada paksaan dari pihak mana pun.

Tangerang Selatan, 20 Juni 2025



Muhammad Rayhan Athaurrahman

**PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Rayhan Athaurrahman
NIM : 2110512160
Program Studi : S1 Sistem Informasi
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa saya memberikan izin sepenuhnya kepada Universitas Pembangunan Nasional
"Veteran" Jakarta untuk memublikasikan karya ilmiah saya yang berjudul:

**"Penerapan Klasifikasi Proses Seleksi Penerimaan Mahasiswa Magang Kampus Merdeka
di PT. XYZ"**

untuk kepentingan akademis, pengembangan ilmu pengetahuan, dan referensi bagi civitas
akademika, tanpa mendapatkan imbalan dalam bentuk apa pun. Saya menyadari bahwa hak cipta
karya ilmiah ini tetap berada pada saya sebagai penulis dan Universitas Pembangunan Nasional
"Veteran" Jakarta.

Saya juga menyatakan bahwa karya ilmiah ini adalah asli hasil karya saya sendiri, bukan hasil
plagiarisme, dan belum pernah dipublikasikan dalam bentuk apa pun sebelumnya, kecuali
disebutkan sumbernya secara jelas.

Demikian pernyataan ini saya buat dengan sebenar-benarnya dan tanpa ada paksaan dari pihak
mana pun.

Jakarta, Juli 2025


Muhammad Rayhan Athaurrahman
2110512160

LEMBAR PENGESAHAN

Judul : Penerapan Klasifikasi Proses Seleksi Penerimaan Mahasiswa
Magang Kampus Merdeka Di PT. XYZ
Nama : Muhammad Rayhan Athaurrahman
NIM : 2110512160

Disetujui oleh:

Penguji 1 :
Iin Ernawati, S.Kom., M.Si.

Penguji 2 :
Novi Trisman Hadi, S.Pd., M.Kom.

Pembimbing 1 :
Ika Nurlaili Isnainiyah, S.Kom., M.Sc.

Pembimbing 2 :
M. Octaviano, S.Kom., M.Kom

Diketahui oleh:

Koordinator Program Studi:
Anita Muliawati, S.Kom., MTL
NIP. 197005212021212002

Dekan Fakultas Ilmu Komputer:
Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM
NIP. 197605082003121002

Tanggal Ujian Tugas Akhir :

15 Mei 2025



Abstrak

Program magang MSIB Kampus Merdeka memberikan pengalaman kerja nyata kepada mahasiswa melalui kolaborasi dengan perusahaan, seperti PT. XYZ, guna meningkatkan kompetensi dan kesiapan mereka menghadapi dunia profesional. Namun, tingginya jumlah pelamar magang menimbulkan tantangan dalam proses seleksi, sehingga diperlukan sistem rekrutmen yang efisien, objektif, dan mampu menyaring kandidat terbaik secara adil. Penelitian ini bertujuan untuk mengimplementasikan model klasifikasi dalam proses administrasi rekrutmen mahasiswa magang di program Kampus Merdeka di PT. XYZ. Fokus utama penelitian ini adalah untuk mengeksplorasi teknik yang digunakan dalam implementasi model klasifikasi dan mengevaluasi hasil pengujian yang dilakukan. Metodologi yang digunakan meliputi penggabungan data dari berbagai sumber, pemrosesan data, dan penerapan algoritma klasifikasi menggunakan Categorical Naive Bayes. Hasil penelitian menunjukkan bahwa model Naive Bayes tanpa SMOTE menghasilkan akurasi 89% namun dengan F1-Score untuk kelas diterima hanya 33%. Fitur SMOTE mencegah terjadinya *data imbalance* dikelas minoritas sehingga model yang dibuat dapat menghasilkan akurasi 80% dan dengan F1-Score untuk kelas diterima 82%. Penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam meningkatkan efisiensi dan efektivitas proses rekrutmen di PT. XYZ.

Kata Kunci: *Predictive Analytics*, Rekrutmen, *Categorical Naive Bayes*, SMOTE, Python

Abstract

The MSIB Kampus Merdeka internship program provides students with real work experience through collaboration with companies, such as PT. XYZ, in order to enhance their competencies and readiness for the professional world. However, the high number of internship applicants presents challenges in the selection process, necessitating a recruitment system that is efficient, objective, and capable of fairly selecting the best candidates. This study aims to implement a classification model in the administrative recruitment process for internship students in the Kampus Merdeka program at PT. XYZ. The main focus of this research is to explore the techniques used in the implementation of the classification model and evaluate the results of the testing conducted. The methodology includes data integration from various sources, data processing, and the application of a classification algorithm using Categorical Naive Bayes. The results show that the Naive Bayes model without SMOTE achieved an accuracy of 89%, but the F1-score for the accepted class was only 33%. The SMOTE feature prevents data imbalance in the minority class, enabling the model to achieve 80% accuracy and an F1-score of 82% for the accepted class. This research is expected to contribute significantly to improving the efficiency and effectiveness of the recruitment process at PT. XYZ.

Keywords: Predictive Analytics, Recruitment, Categorical Naive Bayes, SMOTE, Python

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi ini yang berjudul "PENERAPAN KLASIFIKASI PROSES SELEKSI PENERIMAAN MAHASISWA MAGANG KAMPUS MERDEKA DI PT. XYZ", untuk memperoleh gelar Sarjana di Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Penyusunan skripsi ini tidak terlepas dari bantuan, dukungan, dan bimbingan yang diberikan oleh berbagai pihak. Oleh karena itu, pada kesempatan ini, penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT. yang telah memberikan nikmat dan anugerahnya kepada saya sehingga dapat menyelesaikan penelitian ini, serta nabi-nya Muhammad Salallahu Alaihi Wasallam yang telah menyampaikan kabar gembira dan peringatan bagi seluruh umat manusia. Semoga Allah dapat memberikan nikmat dan perlindungannya kepada seluruh hamba-hamba nya;
2. Keluarga tercinta yang selalu memberikan doa, dukungan, dan motivasi kepada penulis dalam setiap langkah perjalanan studi ini;
3. Bapak Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM, selaku Dekan Fakultas Ilmu Komputer, dan Ibu Anita Anita Muliawati, S.Kom., MTI selaku Koordinator Program Studi S1 Sistem Informasi serta seluruh jajaran dosen dan karyawan Universitas Pembangunan Nasional Veteran Jakarta, atas kesempatan dan fasilitas yang diberikan selama masa studi;
4. Ibu Ika Nurlaili Isnainiyah, S.Kom, M.Sc, Bapak M. Octaviano, S.Kom.,M.Kom, Ibu Iin Ernawati, S.Kom., M.Si., serta Bapak Novi Trisman Hadi, S.Pd., M.Kom. selaku pembimbing dan penguji skripsi yang telah memberikan arahan, bimbingan, dan motivasi selama proses penelitian ini;
5. Mentor Imam B. serta seluruh jajaran karyawan PT. XYZ yang telah memberikan kesempatan magang, izin, dan bantuan dalam pengumpulan data serta informasi yang diperlukan untuk penelitian ini.
6. Teman-teman dan rekan-rekan baik dari lingkungan UPN Veteran Jakarta maupun diluar yang telah membantu baik dukungan moral, materil, atau yang lainnya.
7. Serta semua elemen pendukung yang saya tidak dapat sebutkan satu per satu.

Saran dan kritik yang membangun sangat diharapkan untuk perbaikan di masa yang akan datang, karena penulis memahami bahwa Skripsi ini masih jauh dari kesempurnaan. Semoga skripsi ini dapat memberikan manfaat dan kontribusi yang positif, baik bagi pengembangan ilmu pengetahuan maupun bagi PT. XYZ dan industri terkait.

Terima kasih.

Muhammad Rayhan Athaurrahman
Jakarta Selatan, 1 April 2025

DAFTAR ISI

LEMBAR PERNYATAAN ORISINALITAS	iii
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	iv
LEMBAR PENGESAHAN	v
KATA PENGANTAR	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR RUMUS.....	xiii
PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Penelitian.....	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Luaran Yang Diharapkan	5
1.7 Sistematika Penulisan	5
TINJAUAN PUSTAKA.....	7
2.1 Sistem Informasi	7
2.2 Data Mining.....	7
2.3 Predictive Analytics.....	8
2.4 Data Preprocessing.....	8
2.5 Metode Naïve Bayes	8
2.6 SMOTE	9
2.7 Evaluasi Model	9
2.7.1. Confusion Matrix.....	9
2.7.2. Accuracy	10
2.7.3. Precision.....	10
2.7.4. Recall	10
2.7.5. F ₁ Score.....	10
2.8 Python	11
2.9 Google Colab.....	11
2.10 Penelitian Terdahulu	11

METODE PENELITIAN.....	14
3.1 Alur penelitian.....	14
3.2 Identifikasi Masalah.....	14
3.3 Studi Literatur.....	15
3.4 Pengumpulan Dataset.....	15
3.5 Data Preprocessing.....	15
3.6 SMOTE	15
3.7 Data Modelling	15
3.8 Pengujian Model.....	15
3.9 Implementasi Model	16
3.10 Deployment	16
3.11 Waktu dan Tempat Penelitian	16
3.12 Jadwal Penelitian	16
3.13 Perangkat Penelitian.....	17
HASIL DAN PEMBAHASAN	18
4.3. Pengumpulan Dataset.....	22
4.4. Data Pre-processing	25
4.5. Data Modeling	38
4.6. Pengujian Model.....	39
4.7. Implementasi Prediksi.....	49
PENUTUP.....	54
5.1. Kesimpulan.....	54
5.2. Saran.....	54
DAFTAR PUSTAKA	56

DAFTAR TABEL

Tabel 1 Dataset Pengujian	22
Tabel 2 Tipe Dataset Pengujian	23
Tabel 3 Sampel data kandidat	25
Tabel 4 Sampel data setelah dilakukan penggabungan	27
Tabel 5 Sampel data setelah dilakukan rename kolom	29
Tabel 6 Sampel data setelah dilakukan penggabungan dan rename kolom.....	31
Tabel 7 Sampel data dan hasil pengkategorian IPK.....	31
Tabel 8 Sampel data dan hasil pengkategorian placement	33
Tabel 9 Sampel data dan hasil pengkategorian universitas	34
Tabel 10 Sampel data dan hasil pengkategorian organisasi	35
Tabel 11 Sampel data dan hasil pengkategorian kecocokan jurusan	36
Tabel 12 Sampel data setelah dilakukan label encoding	37
Tabel 13 Hasil evaluasi model Non-SMOTE	38
Tabel 14 Confusion Matrix model Non-SMOTE	38
Tabel 15 Hasil evaluasi model SMOTE	41
Tabel 16 Confusion Matrix model SMOTE	41
Tabel 17 Hasil evaluasi model SMOTE	43
Tabel 18 Label dan hasil prediksi dari kolom ipk_class	44
Tabel 19 Label dan hasil prediksi dari kolom placement_class.....	44
Tabel 20 Label dan hasil prediksi dari kolom univ_class	45
Tabel 21 Label dan hasil prediksi dari kolom org_class	46
Tabel 22 Label dan hasil prediksi dari kolom major_match_class	46

DAFTAR GAMBAR

Gambar 1 Alur Penelitian	14
Gambar 2 Diagram konteks sistem rekrutmen HCM PT XYZ.....	19
Gambar 3 DFD Level 0 sistem rekrutmen magang HCM PT XYZ.....	19
Gambar 4 DFD Level 1 sistem rekrutmen magang HCM PT XYZ.....	20
Gambar 5 DFD Level 2 sistem rekrutmen magang HCM PT XYZ.....	21
Gambar 6 Import library python	25
Gambar 7 Penggabungan File Excel dan Normalisasi Nama Kolom	26
Gambar 8 Pembersihan String, Penghapusan Duplikat, dan Penggabungan Dataframe	27
Gambar 9 Penamaan Kolom yang Deskriptif dan Ekspor Dataframe ke File Excel	28
Gambar 10 Penilaian Berbasis IPK dan Aspek Lain.....	30
Gambar 11 Klasifikasi IPK Mahasiswa Berdasarkan Rentang Nilai	31
Gambar 12 Penilaian Kesesuaian Lokasi antara Kampus dan Tempat Penempatan	32
Gambar 13 Penilaian Universitas Berdasarkan Status Kemitraan dan Peringkat.....	34
Gambar 14 Penilaian Aktivitas Organisasi Berdasarkan Jumlah Sertifikat	35
Gambar 15 Penilaian Kecocokan Jurusan antara Calon dan Jurusan Penempatan.....	36
Gambar 16 Proses Encoding Label dan Penanganan Ketidakseimbangan dengan SMOTE	37
Gambar 17 Pemilihan fitur, target, SMOTE, train_test_split, dan pemilihan metode pemodelan	38
Gambar 18 Evaluasi Model Naive Bayes, Visualisasi Hasil Prediksi, dan Penyimpanan Model Terlatih	38
Gambar 19 Grafik distribusi model Non-SMOTE.....	40
Gambar 20 Grafik distribusi model SMOTE.....	41
Gambar 21 Perbandingan data Non-SMOTE dan SMOTE (train_test_size 30%)	43
Gambar 22 Tree (max_depth=5) dari rules yang telah dibuat	47
Gambar 23 Tree (dalam bentuk if-else) dari rules yang telah dibuat.....	49
Gambar 24 Antarmuka Pengguna Aplikasi TroopersSelect	50
Gambar 25 Struktur Kelas `MainWindow` sebagai Antarmuka Pengguna Aplikasi dengan Inisialisasi dan Penanganan Aksi.....	51
Gambar 26 Implementasi Metode `select_folder` dan `execute_process` untuk Interaksi Pengguna dan Pemrosesan Data	52

DAFTAR RUMUS

2.1 Metode Naive Bayes.....	8
2.2 Accuracy	10
2.3 Precision.....	10
2.4 Recall	10
2.5 F1-Score	10