

PROPOSAL TUGAS AKHIR

“Analisis Komparatif Kinerja Klasifikasi *Support Vector Machine (SVM)* dan Model *Baseline* pada Prediksi Nilai Akhir Mahasiswa Perguruan Tinggi”



Disusun oleh :

Rahma Fitria Tunnisa : 241572010009

Rahmawati : 241572010012

Dosen Pengampu : Hendri Kharisma S.Kom, M.T

PROGRAM STUDI SISTEM INFORMASI

SEKOLAH TINGGI MANAJEMEN INFORMATIKA KOMPUTER

Jl. Raya Dramaga Blok Radar Baru No.8, RT.03/RW.03, Margajaya, Kec. Bogor Bar., Kota Bogor, Jawa Barat 16116, Indonesia

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Kinerja akademik mahasiswa merupakan indikator vital bagi keberhasilan institusi pendidikan tinggi. Dengan volume data akademik, demografi, dan kebiasaan belajar yang terus meningkat, bidang *Educational Data Mining* (EDM) menawarkan peluang besar untuk memprediksi hasil akhir mahasiswa. Prediksi yang akurat memungkinkan pengembangan **Sistem Peringatan Dini (Early Warning System)**, yang memungkinkan dosen dan penasihat akademik memberikan intervensi tepat waktu kepada mahasiswa yang berisiko. Penelitian ini berfokus pada analisis komparatif, menguji efektivitas model klasifikasi canggih **Support Vector Machine (SVM)** terhadap model *baseline* yang telah dipublikasikan pada dataset kinerja mahasiswa yang sama.

1.2. Rumusan Masalah

Berdasarkan latar belakang tersebut, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana membangun dan mengoptimasi model klasifikasi **Support Vector Machine (SVM)** untuk memprediksi kategori Nilai Akhir Mahasiswa (*Output Grade*) pada dataset *Higher Education Students Performance Evaluation?*
2. Bagaimana kinerja klasifikasi model SVM dibandingkan secara komparatif dengan hasil model *baseline* (**KNN, Random Forest, ANN**) yang telah ada pada dataset yang sama?

1.3. Tujuan Penelitian

1. Mengembangkan model klasifikasi SVM yang teroptimasi (*hyperparameter-tuned*) untuk prediksi Nilai Akhir Mahasiswa.
2. Melakukan analisis perbandingan yang valid antara kinerja SVM dengan model *baseline* (KNN, Random Forest, ANN) pada dataset yang sama.
3. Menarik kesimpulan mengenai efektivitas SVM sebagai metode klasifikasi prediktif dalam kasus kinerja akademik.

1.4. Manfaat Penelitian

- **Bagi Akademisi dan Peneliti:** Memberikan wawasan ilmiah mengenai penerapan dan kinerja algoritma SVM pada dataset kinerja mahasiswa, serta menjadi referensi perbandingan yang baru.
- **Bagi Institusi Pendidikan (STMIK TAZKIA):** Menyediakan dasar metodologis untuk mengidentifikasi mahasiswa yang berisiko rendah atau tinggi terhadap pencapaian akademik, mendukung intervensi dini dan peningkatan kualitas pembelajaran.

BAB II

TINJAUAN DATA DAN REFERENSI

2.1. Deskripsi Dataset

| Komponen | Detail |
|---------------------|--|
| Nama Dataset | Higher Education Students Performance Evaluation |
| Sumber Data | Data survei dan akademik dari 145 Mahasiswa di Siprus (2019) |
| Jumlah Sampel | 145 Instances |
| Variabel Input (X) | 31 variabel fitur (termasuk demografi, keluarga, kebiasaan belajar) |
| Variabel Target (Y) | Output Grade (Klasifikasi Multikelas: Fail, DD, DC, CC, CB, BB, BA, AA) |

Link dataset

[Higher Education Students Performance Evaluation - UCI Machine Learning Repository](https://archive.ics.uci.edu/ml/datasets/Higher+Education+Students+Performance+Evaluation)

2.2. Model Referensi

- **Judul Paper:** A Study of Factors Affecting Learning Efficiency on Higher Education Student Performance Evaluation Dataset Using Feature Selection Techniques
- **Penulis:** Hengpraprophm et al.
- **Tahun:** 2022
- **Model Baseline:** K-Nearest Neighbor (KNN), Random Forest, Artificial Neural Network (ANN), dan Linear Regression. Hasil kinerja dari model-model ini akan menjadi *benchmark* utama.

BAB III

METODOLOGI

3.1. Pra-pemrosesan Data

- Membersihkan data (*data cleaning*) dan penanganan nilai nol (*null values*) serta duplikat.
- Mengubah data kategorik ke numerik (*label encoding* atau *one-hot encoding*).
- Melakukan **Standarisasi/Normalisasi** data untuk mengoptimalkan kinerja model SVM.

3.2. Modeling dan Eksperimen

- Menggunakan algoritma klasifikasi utama:
 - **Support Vector Machine (SVM)** sebagai model baru yang diusulkan.
 - **Model Baseline:** K-Nearest Neighbor (KNN), Random Forest, dan Artificial Neural Network (ANN)
 - **Tuning Hyperparameter:** Melakukan optimasi parameter model SVM menggunakan teknik *Cross-Validation*.
- **Tools:** Python (pandas, scikit-learn, seaborn).

3.3. Evaluasi Kinerja dan Visualisasi

- **Confusion Matrix.**
- Metrik Kinerja Klasifikasi: **Akurasi, Presisi, Recall, dan F1-Score.**
- Visualisasi Perbandingan Kinerja: Grafik batang (*bar chart*) yang membandingkan metrik **Accuracy** model SVM dengan Akurasi tertinggi model *baseline* dari *paper* referensi.

BAB IV

KESIMPULAN DAN RENCANA

- Penelitian ini bertujuan untuk membangun model prediksi performa akademik mahasiswa berbasis data **Higher Education Students Performance Evaluation**. Dengan metode klasifikasi SVM, model diharapkan dapat memberikan gambaran yang akurat mengenai kinerja mahasiswa untuk keperluan Sistem Peringatan Dini Akademik.
- Langkah selanjutnya adalah:
 - Melakukan eksperimen *modeling* model SVM dan *tuning hyperparameter* secara mendalam.
 - Visualisasi hasil model dan penghitungan metrik kinerja.
 - Melakukan analisis komparatif yang ketat terhadap hasil SVM versus Model *Baseline* (berdasarkan *paper* referensi).
 - Penyusunan laporan akhir *modeling* disertai grafik dan analisis hasil.