

LAPORAN PENELITIAN

**Mengatasi Ketidakseimbangan Kelas Dalam Pembelajaran Mesin Untuk Memprediksi
Kelulusan Mahasiswa Tepat Waktu di Universitas Islam Riau**



AKMAR EFENDI

225430038

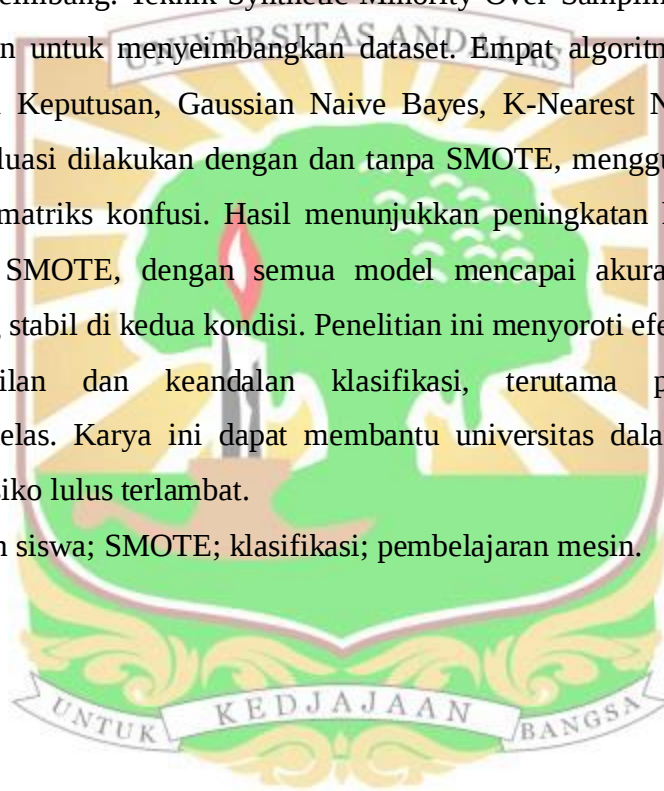
**PROGRAM STUDI PENDIDIKAN PROFESI INSINYUR
SEKOLAH PASCA SARJANA UNIVERSITAS ANDALAS**

2025

Abstrak

Kelulusan tepat waktu merupakan indikator penting kinerja akademik di pendidikan tinggi. Namun, masih banyak mahasiswa yang gagal lulus tepat waktu, sehingga mendorong perlunya model prediktif untuk mendukung pengambilan keputusan akademik. Penelitian ini bertujuan untuk menganalisis dampak ketidakseimbangan kelas terhadap kinerja algoritma pembelajaran mesin dalam memprediksi kelulusan mahasiswa di Universitas Islam Riau. Data diperoleh melalui kuesioner dan diberi label menjadi kelas "lulus tepat waktu" dan "tidak tepat waktu", yang awalnya tidak seimbang. Teknik Synthetic Minority Over-Sampling (SMOTE) diterapkan selama prapemrosesan untuk menyeimbangkan dataset. Empat algoritma pembelajaran mesin dibandingkan: Pohon Keputusan, Gaussian Naive Bayes, K-Nearest Neighbors, dan Support Vector Machine. Evaluasi dilakukan dengan dan tanpa SMOTE, menggunakan akurasi, presisi, recall, skor F1, dan matriks konfusi. Hasil menunjukkan peningkatan kinerja yang signifikan setelah menerapkan SMOTE, dengan semua model mencapai akurasi sekitar 99%. SVM mencapai hasil paling stabil di kedua kondisi. Penelitian ini menyoroti efektivitas SMOTE dalam meningkatkan keadilan dan keandalan klasifikasi, terutama pada dataset dengan ketidakseimbangan kelas. Karya ini dapat membantu universitas dalam intervensi dini bagi mahasiswa yang berisiko lulus terlambat.

Kata Kunci :kelulusan siswa; SMOTE; klasifikasi; pembelajaran mesin.



Abstract

Timely graduation is an important indicator of academic performance in higher education. However, many students still fail to graduate on time, prompting the need for predictive models

to support academic decision-making. This study aims to analyze the impact of class imbalance on machine learning algorithm performance in predicting student graduation at the Islamic University of Riau. Data were obtained through questionnaires and labeled into “graduated on time” and “not on time” classes, which were initially imbalanced. The Synthetic Minority Over-Sampling Technique (SMOTE) was applied during preprocessing to balance the dataset. Four machine learning algorithms were compared: Decision Tree, Gaussian Naive Bayes, K-Nearest Neighbors, and Support Vector Machine. The evaluation was conducted with and without SMOTE, using accuracy, precision, recall, F1-score, and confusion matrix. Results showed significant performance improvements after applying SMOTE, with all models achieving around 99% accuracy. SVM achieved the most stable results across both conditions. The study highlights the effectiveness of SMOTE in improving classification fairness and reliability, especially in datasets with class imbalance. This work may assist universities in early intervention for students at risk of late graduation.

Keywords : student graduation; SMOTE; classification; machine learning.

