

KLASIFIKASI MAHASISWA BERPOTENSI *DROP OUT* (DO) MENGGUNAKAN ALGORITMA *RANDOM FOREST*

Nama Mahasiswa : Mitha Zalina Ramadhany
Nim : 6304221424
Dosen Pembimbing : Lidya Wati, M. Kom

ABSTRAK

Mahasiswa *drop out* merupakan masalah serius yang berdampak pada penurunan kualitas lulusan dan reputasi perguruan tinggi. *Drop out* mahasiswa menjadi tantangan serius di perguruan tinggi, termasuk Politeknik Negeri Bengkalis yang mencatat 198 kasus *drop out*. Penelitian ini bertujuan membangun sistem klasifikasi mahasiswa berpotensi *drop out* menggunakan algoritma *Random Forest* berbasis data akademik dan non-akademik. Metode *Rapid Application Development* (RAD) digunakan dalam pengembangan sistem, dengan dataset 398 mahasiswa dan sembilan atribut penelitian yaitu Jenis Kelamin, Lama Studi, Pelanggaran, IPK, Ketidakhadiran, Penghasilan Orang Tua, Riwayat Cuti dan Registrasi, Status KIP dan Asal Tempat Tinggal. Hasil evaluasi menunjukkan akurasi 97%, *precision* 98%, dan *recall* 97%. Fitur *early warning* berhasil mengidentifikasi 49 mahasiswa dari 200 mahasiswa aktif (24%) dari total keseluruhan data yang digunakan sebagai berpotensi *drop out*. Hasil ini menunjukkan bahwa algoritma *Random Forest* mampu digunakan sebagai sistem klasifikasi mahasiswa berpotensi *drop out* secara terukur dan berbasis data.

Kata Kunci: *Random Forest, Drop Out, Early Warning, Klasifikasi Mahasiswa, Machine Learning.*

CLASSIFICATION OF STUDENTS AT RISK OF DROPPING OUT (DO) USING THE RANDOM FOREST ALGORITHM

Nama Mahasiswa : Mitha Zalina Ramadhany
Nim : 6304221424
Dosen Pembimbing : Lidya Wati, M. Kom

ABSTRACT

Student dropout is a serious problem that affects the quality of graduates and the reputation of higher education institutions. Student dropout is a serious challenge in higher education institutions, including Bengkalis State Polytechnic, which recorded 198 cases of dropout. This study aims to develop a classification system for students at risk of dropping out using the Random Forest algorithm based on academic and non-academic data. The Rapid Application Development (RAD) method was used in the system development, with a dataset of 398 students and nine research attributes, namely Gender, Length of Study, Violations, GPA, Absence, Parental Income, Leave and Registration History, KIP Status, and Place of Residence. The evaluation results showed an accuracy of 97%, precision of 98%, and recall of 97%. The early warning feature successfully identified 49 students out of 200 active students (24%) from the total data used as potentially dropping out. These results indicate that the Random Forest algorithm can be used as a measurable and data-based classification system for students who are potentially dropping out.

Keywords: *Random Forest, Drop Out, Early Warning, Student Classification, Machine Learning.*