

RINGKASAN

Drop out mahasiswa merupakan salah satu permasalahan yang dapat memengaruhi mutu pendidikan dan tingkat kelulusan di perguruan tinggi. Oleh karena itu, diperlukan suatu model klasifikasi yang mampu memprediksi mahasiswa yang berpotensi mengalami drop out. Penelitian ini bertujuan untuk membangun serta membandingkan performa algoritma Decision Tree C5.0 dengan dan tanpa penerapan seleksi fitur Mutual Information dalam klasifikasi mahasiswa berisiko drop out. Data penelitian mencakup mahasiswa S1 angkatan 2018–2023 dengan atribut performa akademik setiap semester serta aktivitas tugas akhir mahasiswa. Proses pengolahan data dilakukan melalui beberapa tahapan, yaitu seleksi data, pembersihan, pelabelan, transformasi, dan penanganan ketidakseimbangan kelas menggunakan metode Random Undersampling. Evaluasi model dilakukan dengan menggunakan metrik akurasi, presisi, recall, dan f1-score. Hasil penelitian menunjukkan bahwa pada mahasiswa angkatan 2022–2023, model tanpa seleksi fitur menghasilkan akurasi sebesar 87,14% dan meningkat menjadi 88,57% setelah penerapan seleksi fitur. Sementara itu, pada mahasiswa angkatan 2018–2021, akurasi model tanpa seleksi fitur mencapai 93,04% dan meningkat menjadi 94,30% setelah penerapan seleksi fitur. Faktor dominan yang memengaruhi klasifikasi berbeda pada masing-masing kelompok, di mana jumlah absen menjadi indikator utama pada mahasiswa angkatan 2022–2023, sedangkan durasi pengerjaan tugas akhir lebih berpengaruh pada mahasiswa angkatan 2018–2021. Berdasarkan hasil tersebut, penelitian ini menunjukkan bahwa penerapan seleksi fitur dapat meningkatkan performa model klasifikasi C5.0. Selain itu, model yang dibangun diharapkan dapat membantu pihak universitas dalam mengidentifikasi mahasiswa berisiko drop out lebih awal, sehingga strategi pencegahan dan pendampingan akademik dapat dilakukan secara lebih tepat sasaran.