

1. Sitasi Artikel

Wijaya, A., & Nugroho, L. E. (2022). Sistem Prediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma C4.5 Berbasis K-Means Clustering. JOINTECS, 7(1), 45-52.

2. Latar Belakang & Tujuan

Tingginya angka putus sekolah (drop out) menjadi masalah serius bagi institusi pendidikan tinggi. Penelitian ini penting karena bertujuan mengembangkan model prediksi untuk mengidentifikasi mahasiswa berisiko sejak dini, sehingga intervensi dapat dilakukan. Tujuan utamanya adalah membangun sistem prediksi kelulusan mahasiswa menggunakan kombinasi algoritma K-Means untuk pengelompokan data awal dan C4.5 untuk klasifikasi.

3. Metode

Penelitian ini menggunakan pendekatan kuantitatif dengan teknik data mining. Data yang digunakan adalah data akademik historis mahasiswa dari tahun 2015 hingga 2019, mencakup variabel seperti IPK semester, status ekonomi, dan keaktifan organisasi. Prosesnya diawali dengan preprocessing data, kemudian data dikelompokkan menjadi tiga cluster (berisiko tinggi, sedang, rendah) menggunakan K-Means. Hasil clustering ini kemudian menjadi atribut baru yang digunakan bersama atribut lainnya untuk membangun pohon keputusan dengan algoritma C4.5 guna mengklasifikasikan status kelulusan (Tepat Waktu, Terlambat, Drop Out).

4. Hasil/Temuan Kunci

Model klasifikasi yang diusulkan berhasil mencapai akurasi sebesar 89.5%. Pengujian menunjukkan bahwa atribut hasil clustering K-Means dan IPK semester 1-4 merupakan faktor paling berpengaruh dalam menentukan prediksi kelulusan. Sistem mampu mengidentifikasi 92% mahasiswa yang masuk dalam kategori berisiko tinggi secara akurat, membuktikan efektivitas model untuk deteksi dini.

5. Kontribusi & Keterbatasan

Kontribusi utama penelitian ini adalah penerapan model hibrida K-Means dan C4.5 yang terbukti efektif untuk prediksi kelulusan di konteks data akademik lokal. Namun, penelitian ini memiliki keterbatasan karena hanya menggunakan data dari satu institusi, sehingga generalisasi model ke institusi lain memerlukan pengujian lebih lanjut. Selain itu, faktor non-akademik seperti motivasi personal belum diikutsertakan sebagai variabel.

6. Takeaway Anda

Pelajaran utama yang didapat adalah bagaimana teknik data mining dapat memberikan solusi praktis untuk masalah institusional. Kombinasi metode unsupervised dan supervised learning ternyata dapat meningkatkan performa model prediksi secara signifikan.