

ABSTRAK

Pemilihan peminatan dalam studi Informatika merupakan proses penting yang menentukan arah akademik dan karier mahasiswa. Namun, pengambilan keputusan sering kali tidak didasarkan pada data akademik yang objektif. Penelitian ini bertujuan mengembangkan sistem rekomendasi peminatan berbasis machine learning dengan memanfaatkan algoritma Principal Component Analysis (PCA) dan Decision Tree Classifier. Data yang digunakan berupa nilai akademik mahasiswa program studi Informatika yang telah dikonversi dari huruf ke angka. Proses awal mencakup pembersihan data (menghapus kolom tidak relevan seperti nama dan semester), imputasi nilai kosong dengan rata-rata kolom, serta penanganan ketidakseimbangan data menggunakan teknik undersampling. Reduksi dimensi dilakukan menggunakan PCA, menghasilkan 9 komponen utama yang merepresentasikan 87% variasi data. Komponen ini digunakan sebagai fitur input ke model Decision Tree Classifier. Model dilatih dengan pembagian data 80% pelatihan dan 20% pengujian, serta dievaluasi menggunakan confusion matrix, accuracy score, dan classification report. Hasil awal menunjukkan kinerja klasifikasi yang baik, kemudian ditingkatkan melalui hyperparameter tuning menggunakan GridSearchCV dengan cross-validation (cv=5). Parameter terbaik diperoleh pada `n_components=9`, `max_depth=3`, `min_samples_split=2`, dan `min_samples_leaf=5`, menghasilkan akurasi pengujian sebesar 83,33%, dengan penjelasan variasi sebesar 89,50% dan rata-rata akurasi cross-validation sebesar $46\% \pm 26,53\%$. Model terbaik ini kemudian diimplementasikan ke dalam sistem rekomendasi berbasis GUI menggunakan Python Tkinter, yang memungkinkan mahasiswa memperoleh rekomendasi peminatan secara langsung berdasarkan input nilai akademik mereka.

Kata Kunci: akurasi, Decision Tree, data akademik, Machine Learning, Principal Component Analysis.