

ABSTRAK

Penentuan kategori Uang Kuliah Tunggal (UKT) yang masih dilakukan secara manual berpotensi menimbulkan subjektivitas dan ketidakkonsistenan dalam pengambilan keputusan. Penelitian ini mengusulkan pendekatan hybrid berbasis data mining dengan mengintegrasikan Principal Component Analysis (PCA), K-Means Clustering, dan Naïve Bayes Classifier untuk menghasilkan sistem klasifikasi UKT yang lebih objektif dan sistematis. PCA digunakan untuk mereduksi dimensi dan menghilangkan redundansi fitur, K-Means untuk membentuk pseudo-label berbasis kemiripan karakteristik sosial ekonomi, serta Naïve Bayes untuk memodelkan proses klasifikasi. Dataset yang digunakan terdiri dari 452 data mahasiswa dengan atribut sosial ekonomi. Hasil reduksi dimensi menunjukkan bahwa 18 fitur dapat direpresentasikan menjadi 12 komponen utama dengan mempertahankan 95% variasi data. Penentuan jumlah cluster optimal menghasilkan $k=8$ pada skenario PCA dan $k=7$ tanpa PCA, yang menunjukkan bahwa reduksi dimensi mempengaruhi struktur data. Evaluasi model menunjukkan performa akurasi tinggi pada seluruh algoritma, namun analisis lebih lanjut mengindikasikan bahwa model berbasis PCA memiliki stabilitas dan kemampuan generalisasi yang lebih baik, ditunjukkan oleh konsistensi hasil cross-validation dan selisih kecil antara data latih dan data uji. Fokus pada metrik recall menunjukkan bahwa model mampu mengidentifikasi mayoritas kelas secara efektif. Secara keseluruhan, pendekatan ini menghasilkan model yang lebih robust dan layak digunakan sebagai pendukung keputusan dalam penentuan UKT yang lebih adil dan konsisten.

Kata Kunci: Uang Kuliah Tunggal (UKT), Principal Component Analysis (PCA), K-Means Clustering, Naïve Bayes Classifier, Machine Learning