

ABSTRAK

Penelitian ini menganalisis distribusi nilai Tes Wawasan Kebangsaan (TWK), Tes Intelegensi Umum (TIU), dan Tes Karakteristik Pribadi (TKP) dari 2.490 peserta tes CPNS Kemenkeu 2024 menggunakan *machine learning* untuk mengidentifikasi kelemahan dalam desain tes dan merekomendasikan perbaikan berbasis data. *Clustering* dengan metode *K-Means* ($k=3$) mengelompokkan peserta ke dalam tiga klaster: TKP tinggi namun TWK/TIU rendah (65% lulus), nilai seimbang (55% lulus), dan TKP rendah namun TWK tinggi (30% lulus), dengan nilai *Silhouette Score* sebesar 0,59, yang menunjukkan kualitas klaster yang cukup baik. Klasifikasi menggunakan *Logistic Regression* (akurasi 0,7897), *Random Forest* (akurasi 0,9779), dan *XGBoost* (akurasi 0,9809) menunjukkan performa sangat baik, terutama pada dua model ensemble. Namun, tingginya tingkat *False Negative* mencerminkan dominasi nilai TKP yang kurang diskriminatif serta rendahnya variasi pada nilai TIU dan TWK, sehingga model masih menghadapi tantangan dalam memprediksi peserta yang lulus dengan tepat. Temuan ini mengindikasikan adanya ketidakseimbangan bobot antar tes, sehingga diperlukan revisi pada soal TKP dengan penambahan skenario etika yang lebih kompleks, TIU dengan logika dan analisis data yang lebih menantang, serta TWK dengan materi yang lebih relevan terhadap isu kebangsaan kontemporer. Perbaikan ini diharapkan dapat meningkatkan kemampuan diskriminatif tes dan validitas seleksi. Penelitian ini memberikan wawasan empiris untuk mengoptimalkan desain tes CPNS, guna memastikan proses seleksi yang lebih valid, adil, dan berbasis data melalui analisis pola nilai menggunakan metode *machine learning*.

Kata kunci: Prediksi Rekrutmen, *Machine Learning*, Klasifikasi Keputusan Rekrutmen, Seleksi CPNS