

ABSTRAK

Nama : Fina Sifaul Nufus
NIM : 14230025
Program Studi : Ilmu Komputer
Fakultas : Teknologi Informasi
Jenjang : Strata Dua (S2)
Peminatan : *Software Engineering & Data Science*
Judul Tesis : *Cross-project Defect Prediction* pada Dataset AEEEM
Menggunakan *Hybrid SMOTE–Tomek* dan *Ensemble Learning*

Penelitian ini bertujuan untuk meningkatkan kinerja *Cross-project Defect Prediction* (CPDP) pada dataset AEEEM melalui penerapan pendekatan *hybrid preprocessing* yang mengombinasikan normalisasi, reduksi dimensi menggunakan PCA, penyeimbangan kelas dengan *SMOTE–Tomek*, serta penyesuaian *threshold* keputusan. Eksperimen dilakukan pada lima proyek AEEEM, yaitu EQ, JDT, LC, ML, dan PDE, dengan dua skenario utama, yaitu *single-source* CPDP dan *multi-source* CPDP. Model yang digunakan adalah *Random Forest* dan *Support Vector Machine* (SVM), sedangkan kinerja dievaluasi menggunakan metrik *F1-score* dan AUC. Hasil eksperimen menunjukkan bahwa pendekatan *multi-source* secara umum menghasilkan kinerja yang lebih stabil dibandingkan *single-source*. Selain itu, penerapan *hybrid preprocessing* terbukti mampu meningkatkan *F1-score* secara signifikan dibandingkan *baseline*, terutama pada dataset dengan rasio *defect* yang rendah. *Ablation study* mengonfirmasi bahwa peningkatan kinerja diperoleh dari kombinasi penyeimbangan kelas dan *threshold tuning*, bukan dari satu komponen tunggal.

Kata kunci: AEEEM, *Cross-project Defect Prediction*, *Hybrid Preprocessing*, *Machine Learning*, *SMOTE–Tomek*