

INTISARI

Dropout mahasiswa menjadi permasalahan serius bagi perguruan tinggi karena berdampak pada mahasiswa maupun institusi. Penelitian ini bertujuan membandingkan kinerja lima algoritma machine learning (Decision Tree, Random Forest, XGBoost, LightGBM, dan CatBoost) dalam memprediksi dropout mahasiswa, membandingkan hasil penjelasan SHAP dan LIME, menganalisis stabilitas kedua metode penjelasan tersebut, serta merekomendasikan kombinasi algoritma dan metode penjelasan yang paling layak digunakan. Penelitian menggunakan dataset sekunder Predict Students' Dropout and Academic Success (UCI ID: 697) yang memuat 4.424 rekaman mahasiswa dengan 36 fitur, dianalisis melalui pipeline eksperimen mencakup RandomizedSearchCV, uji McNemar, paired t-test, serta empat skenario pengujian stabilitas penjelasan.

Hasil penelitian menunjukkan CatBoost mencapai ROC-AUC tertinggi (0,9364), sedangkan XGBoost meraih F1-Score dan Recall tertinggi (0,8173 dan 0,7641) sehingga paling sesuai untuk sistem peringatan dini. Seluruh model ensemble unggul signifikan dibanding Decision Tree (McNemar dan paired t-test, $p < 0,05$). Perbandingan SHAP dan LIME pada XGBoost menghasilkan kesepakatan top-10 fitur yang rendah (Jaccard=0,4286), namun fitur akademik dan finansial seperti Curricular units 2nd sem (approved) dan Tuition fees up to date konsisten mendominasi kedua metode. TreeSHAP terbukti deterministik sempurna (Jaccard=1,0000), sedangkan LIME hanya stabil sedang pada level top-10 fitur (Jaccard=0,5557) dan rendah pada peringkat penuh (Spearman=0,1860), menunjukkan sensitivitas LIME terhadap variasi random seed. Meskipun CatBoost unggul pada ROC-AUC, selisihnya terhadap XGBoost tergolong sangat tipis (0,9364 berbanding 0,9302) dan secara statistik tidak signifikan (McNemar dan paired t-test, $p > 0,05$), sedangkan skor stabilitas penjelasan (S) antar-model berbasis TreeSHAP relatif seragam sehingga tidak menjadi faktor pembeda. Keunggulan XGBoost pada F1-Score dan Recall menghasilkan skor kinerja ternormalisasi (P) yang lebih tinggi karena rentang variasi F1-Score antar-model jauh lebih lebar dibanding rentang ROC-AUC yang sangat sempit, sehingga pada kerangka penilaian gabungan $CI=0,6P+0,4S$, XGBoost tetap unggul meskipun tidak mencatatkan ROC-AUC tertinggi.

Kerangka penilaian gabungan yang mengintegrasikan kinerja prediktif dan stabilitas penjelasan ($CI=0,6P+0,4S$) menempatkan XGBoost sebagai algoritma paling direkomendasikan untuk sistem peringatan dini dropout mahasiswa, dengan skor $CI=0,9117$ yang konsisten pada seluruh variasi bobot sensitivitas.

Kata kunci: dropout mahasiswa, machine learning, explainable AI, SHAP, LIME

ABSTRACT

Student dropout is a serious problem for higher education institutions because of its impact on both students and the institution. This study aims to compare the performance of five machine learning algorithms (Decision Tree, Random Forest, XGBoost, LightGBM, and CatBoost) in predicting student dropout, to compare the explanations produced by SHAP and LIME, to analyze the stability of both explanation methods, and to recommend the most suitable algorithm-explanation combination. The study uses the secondary dataset Predict Students' Dropout and Academic Success (UCI ID: 697), containing 4,424 student records with 36 features, analyzed through an experimental pipeline covering RandomizedSearchCV, McNemar's test, paired t-test, and four explanation-stability testing scenarios.

The results show that CatBoost achieved the highest ROC-AUC (0.9364), while XGBoost achieved the highest F1-Score and Recall (0.8173 and 0.7641), making it the most suitable for an early-warning system. All ensemble models significantly outperformed the Decision Tree baseline (McNemar and paired t-test, $p < 0.05$). The comparison of SHAP and LIME on XGBoost showed low top-10 feature agreement (Jaccard=0.4286), although academic and financial features such as Curricular units 2nd sem (approved) and Tuition fees up to date consistently dominated both methods. TreeSHAP proved to be perfectly deterministic (Jaccard=1.0000), while LIME showed only moderate stability at the top-10 feature level (Jaccard=0.5557) and low stability in full ranking (Spearman=0.1860), indicating sensitivity to random seed variation. Although CatBoost achieved the highest ROC-AUC, the margin over XGBoost was very narrow (0.9364 versus 0.9302) and statistically non-significant, while stability scores were nearly uniform across tree-based models; XGBoost's higher F1-Score therefore dominated the normalized performance score and, consequently, the combined index.

The combined assessment framework integrating predictive performance and explanation stability ($CI = 0.6P + 0.4S$) identified XGBoost as the most recommended algorithm for a student dropout early-warning system, with a CI score of 0.9117 that remained consistent across all sensitivity weight variations.

Keywords: student dropout, machine learning, explainable AI, SHAP, LIME