

ABSTRACT

Accurate prediction of patient hospitalization duration is a critical challenge in healthcare resource management, as prolonged hospital stays impose substantial costs on both patients and healthcare institutions. Clinical data stored in Electronic Health Records (EHR) in the form of structured tabular data present unique modeling challenges due to their heterogeneous and nonlinear characteristics. This study aims to develop an accurate and interpretable Length of Stay (LoS) prediction model by integrating the Feature Tokenizer Transformer (FT-Transformer) with SHAP-based explainability. The proposed model was developed and evaluated using the Microsoft LoS dataset containing 100,000 patient admission records. Hyperparameter optimization was performed using Optuna, while SHAP values derived from a baseline model were utilized as a model-aware feature selection mechanism across four feature subsets ($k = 5, 10, 15, 20$, and 25), with the top-15 subset identified as the optimal configuration. The final model achieved an MAE of 0.2202, RMSE of 0.3574, RAE of 0.1159, and RSE of 0.0233, representing a 25.96% reduction in MAE compared with XGBoost on the same dataset. Tolerance-based evaluation further showed that 99.1% of predictions were within ± 1 day of the actual LoS. SHAP interpretability was applied at three levels, namely global, local, and risk group analyses. The results consistently identified neutrophils and bloodureanitro as the most influential predictors, while their relative importance varied across patient complexity levels. These findings demonstrate that integrating FT-Transformer with SHAP-based feature selection and multi-level SHAP explainability provides an accurate, competitive, and interpretable framework for clinical LoS prediction, supporting more informed clinical decision-making and healthcare resource management.

Keywords: *eXplainable Artificial Intelligence, Feature Tokenizer Transformer, Length of Stay Prediction, SHAP, Tabular Data*

ABSTRAK

Prediksi yang akurat terhadap lama rawat inap pasien (*Length of Stay* atau LoS) merupakan tantangan penting dalam pengelolaan sumber daya di bidang kesehatan, karena durasi rawat inap yang panjang dapat meningkatkan biaya bagi pasien maupun fasilitas pelayanan kesehatan. Data klinis yang tersimpan dalam *Electronic Health Records* (EHR) dalam bentuk data tabular terstruktur memiliki karakteristik yang heterogen dan nonlinier sehingga menyulitkan proses pemodelan. Penelitian ini bertujuan mengembangkan model prediksi LoS yang akurat dan dapat diinterpretasikan dengan mengombinasikan *Feature Tokenizer Transformer* (FT-*Transformer*) dan interpretabilitas berbasis SHAP. Model dikembangkan menggunakan Microsoft LoS *Dataset* yang terdiri atas 100.000 data rawat inap pasien. Optimasi *hyperparameter* dilakukan menggunakan Optuna, sedangkan nilai SHAP dari model dasar dimanfaatkan sebagai mekanisme seleksi fitur berbasis model pada empat variasi jumlah fitur ($k = 5, 10, 15, 20$, dan semua fitur), dengan 15 fitur terbaik dipilih sebagai konfigurasi optimal. Model yang dihasilkan memperoleh nilai MAE sebesar 0,2202, RMSE sebesar 0,3574, RAE sebesar 0,1159, dan RSE sebesar 0,0233, serta mampu menurunkan nilai MAE sebesar 25,96% dibandingkan model XGBoost pada dataset yang sama. Hasil evaluasi berbasis toleransi menunjukkan bahwa 99,1% prediksi berada dalam rentang ± 1 hari dari nilai LoS aktual. Selain itu, interpretabilitas SHAP diterapkan pada tiga tingkat, yaitu global, lokal, dan kelompok risiko. Hasil analisis secara konsisten menunjukkan bahwa *neutrophils* dan *bloodureanitro* merupakan fitur yang paling berpengaruh terhadap prediksi, dengan tingkat kepentingan relatif yang bervariasi sesuai kompleksitas kondisi pasien. Temuan ini menunjukkan bahwa integrasi FT-*Transformer* dengan seleksi fitur berbasis SHAP dan interpretabilitas SHAP multitingkat menghasilkan kerangka prediksi lama rawat inap yang akurat, kompetitif, dan dapat diinterpretasikan, sehingga mendukung pengambilan keputusan klinis dan pengelolaan sumber daya kesehatan yang lebih baik.

Kata Kunci: Data Tabular, *eXplainable Artificial Intelligence*, *Feature Tokenizer Transformer*, Prediksi Lama Rawat Inap, SHAP