

Analisis Performa dan Interpretasi Model TabTransformer untuk Prediksi Fraud Klaim Asuransi = Performance Analysis and Interpretation of TabTransformer Model for Insurance Claim Fraud Prediction

Larasati Raharjo, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920577640&lokasi=lokal>

Abstrak

Industri asuransi menghadapi tantangan dalam mendeteksi fraud pada klaim yang dapat menyebabkan kerugian finansial dan penurunan kepercayaan masyarakat terhadap perusahaan asuransi. Metode konvensional dinilai kurang efektif karena membutuhkan proses manual dan sumber daya yang besar. Oleh karena itu, pendekatan machine learning diajukan untuk menyelesaikan masalah ini secara efisien. Dalam beberapa tahun terakhir, pendekatan deep learning yang merupakan cabang dari machine learning semakin populer karena kemampuannya menangani data kompleks secara efisien. Salah satu metode deep learning yang dirancang untuk data tabular adalah TabTransformer, yang memanfaatkan mekanisme column embedding dan self-attention untuk menangkap interaksi kompleks antar fitur kategorik. Penelitian ini bertujuan untuk menganalisis performa dan interpretasi dari model TabTransformer dalam mendeteksi fraud pada klaim asuransi. Model TabTransformer dibandingkan dengan model berbasis pohon keputusan seperti Decision Tree, Random Forest, dan XGBoost. Evaluasi dilakukan terhadap tiga dataset klaim asuransi berbeda menggunakan metrik AUC, sensitivity, dan specificity. Hasil penelitian menunjukkan bahwa TabTransformer memiliki kinerja kompetitif dan stabil pada ketiga dataset. Selain itu, analisis interpretabilitas model menunjukkan bahwa fitur-fitur penting yang diidentifikasi oleh model konsisten dengan interpretasi model lainnya, sehingga memperkuat validitas model TabTransformer dalam aplikasi nyata sebagai sistem pendukung keputusan untuk mendeteksi klaim fraud.

.....The insurance industry faces significant challenges in detecting fraud claims, which can lead to financial losses and decreased public trust in insurance companies. Traditional detection methods are often considered ineffective due to their reliance on manual investigation and resource-intensive processes. To address this issue efficiently, machine learning has been proposed as a viable approach. In recent years, deep learning has gained popularity as a subset of machine learning due to its ability to handle complex data efficiently. One of the deep learning methods specifically designed for structured tabular data is the TabTransformer, which utilizes column embeddings and self-attention mechanisms to capture complex interactions between categorical features. This study aims to analyze the performance and interpretability of the TabTransformer model in detecting insurance claim fraud. The model is compared with commonly used tree-based methods, namely Decision Tree, Random Forest, and XGBoost. The evaluation is conducted on three different insurance claim datasets using evaluation metrics such as AUC-ROC, sensitivity, and specificity. Results show that the TabTransformer delivers competitive and stable performance. Furthermore, model interpretability analysis reveals that the most influential features identified are consistent, thereby reinforcing the model's validity as a decision support tool for insurance fraud detection.