Fraudulent Claim Detection

The project addresses the detection of fraudulent of Insurance claims using predictive analytics. Initial steps involve data loading, revealing no null values but reveals a significant class imbalance—with only 6% of claims being fraudulent. Visualizations include a correlation heatmap, identifying strong relationships among variables like Age, DriverRating, and VehiclePrice. A bar plot of claim fraud frequency highlights the class imbalance, critical for modeling strategy. EDA also includes boxplots and count plots, showing fraud is more frequent in claims with lower driver ratings and older vehicles. Preprocessing involves label encoding and scaling, followed by model building using Logistic Regression, Random Forest. F1-score and AUC-ROC, indicating a strong balance between precision and recall. This solution enables early fraud detection, helping insurers reduce financial losses and improve operational efficiency.