Plain XG Boost

Description:

This code implements a loan default prediction model using XGBoost, a powerful gradient boosting algorithm. It preprocesses the dataset by encoding categorical variables into numerical values and standardizing numeric features. The data is split into training and testing sets with stratification to maintain class balance. The model is trained with default XGBoost parameters and evaluated on the test set using accuracy, AUC, and a detailed classification report.

Results:

The model achieved strong performance with an accuracy of 93% and an AUC of 0.94, indicating excellent ability to distinguish between defaulters and non-defaulters. The F1 score for predicting defaulters was 0.83, demonstrating a good balance between precision and recall for the critical risk class. These results establish the model as a reliable baseline for credit risk assessment that can be further improved through hyperparameter tuning, imbalance handling, and explainability techniques.