Predicting Credit Risk Model Stability in Home Credit Using LightGBM Model

(Individual Contribution)

Sanjana K. Y. C.

210572P Grpou-06

1. Model Training:
    - The parameters for the LightGBM model (`lgbm_best_params`) are defined.
    - LightGBM datasets (`lgb_train` and `lgb_val`) are created from the scaled training and validation data.
    - The LightGBM model (`gbm`) is trained using the training dataset with validation monitoring and early stopping.

2. Model Evaluation:
    - The trained model is used to predict probabilities for both the training and validation datasets.
    - The AUC scores are computed for the training and validation datasets using the predicted probabilities.
    - A function `gini_stability` is provided to assess the stability of the model over time. This function calculates a stability score based on the Gini coefficient and some additional parameters.
    - Stability scores are calculated for both the training and validation datasets.

3. Model Prediction:
    - The trained model is used to predict probabilities for the test dataset (`X_test`).
    - The predictions are saved in a CSV file named "submission.csv", along with the corresponding case IDs.