

XGBOOST REGRESSION

This code uses **XGBoost regression** to predict the length of a product based on its product type. Here are the step-by-step explanations of the code:

- The necessary libraries are imported: `pandas` for data manipulation, `xgboost` for XGBoost regression, and `mean_absolute_percentage_error` from `sklearn.metrics` to evaluate the model's performance. The training and testing data are loaded from CSV files.
- The training and testing data are combined into one dataset. Missing values are replaced with 0.
- The training and testing features and target variables are selected: `PRODUCT_TYPE_ID` is the feature and `PRODUCT_LENGTH` is the target.
- An instance of XGBoost regressor is created. The XGBoost model is trained on the training data.
- 8. Predictions are made on the testing data using the trained XGBoost model. The mean absolute percentage error (MAPE) between the actual and predicted target values is calculated.
- The MAPE score is printed. A submission file is created with the predicted target values.

Overall, this code trains an XGBoost regression model on the training data to predict the length of a product based on its product type. The model is then evaluated on the testing data using mean absolute percentage error and the predicted values are saved to a submission file.