

Responsible Machine Learning

Assignment 5

10 points

February 7, 2025

In Assignment 5, you will work with your group to debug your best model following the instructions below. A **template** has been provided with examples of simple sensitivity and residual analysis exercises. Please let me know immediately if you find typos or mistakes in this assignment or related materials.

1 Test How this Lending Model Performs in Recession Conditions.

Cells 7–8 demonstrate a basic stress-testing exercise, in which recession conditions are simulated and model performance is re-estimated under these conditions. Generally, lending model performance degrades quickly when recession conditions arise. You should see that our EBM models are no exception to this rule.

2 Conduct Residual Analysis and Remediate Discovered Bugs.

Cells 9–15 use a basic residual analysis procedure to find outliers and identify a fundamental problem with our data and EBM model. Once these problems are identified, you should be able to increase your model performance by accounting for them. You may use the template approach and/or additional approaches to improve your model.

3 Submit Code Results.

Your deliverable for this assignment is to update your group's GitHub repository to reflect this debugging exercise. Stress-testing is worth 5 points. Remediating your model by removing outliers and handling data imbalance, to increase validation AUC, is worth up to 5 points. Groups with better debugging will receive higher scores.

Your deliverables are due **XX, XX XX^{XX}**, at **11:59 PM ET**.

(HINT: If you perform Assignment 5 correctly, it should allow you to boost your model performance from Assignments 1 and/or 3!)