

MSDS 498: Performance Validation Guide

Getting Started

The final state of the capstone project is to perform the model performance validation that we outlined in the Performance Monitoring Guide. In the first stage of the capstone project we performed

Project Data

The data for this stage of the capstone project the validation data set. Remember that the project data set has three flags on it – train, test, and validate which take values 0/1. Subset this data set to the records where `validate == 1`, and you will have the validation data set.

General Paper Guidelines

The general guidelines are as follows.

- (1) The paper should follow the appropriate final paper format outlined in this guide.
- (2) The paper should be well written. The formats are designed to help you with the overall paper structure. The remainder of the presentation is left to you.
- (3) The term data is the plural of datum. We should write 'the data are', not 'the data is'.
- (4) The body of the paper should be between five and ten pages, double spaced. The title page, the bibliography, and any relevant appendices do not count against the page count.
- (5) Graphics and tables should be labeled so that they can be referenced. Typical labels are 'Figure 1' and 'Table 1'. Graphics and tables should also be large enough to be easily read and centered in the page.
- (6) All papers should have page numbers on each page in the lower right hand corner.

General Grading Guidelines

Your final paper is worth a total of 100 points. Here is how those points will be broken down. Students are expected to follow the instructions for the paper format and content provided in this guide.

- (1) Production Model, Model Development Performance, and Performance Monitoring Plan (50 points)
- (2) Performance Monitoring Results (50 points)

Final Paper Format

Here is an outline for the format of your model development documentation. Students are expected to follow this prescribed format. In addition all papers should contain a title page.

Note that the format of the Performance Validation Results document is almost the same as the Performance Monitoring Plan document.

Title:

Model #101: Credit Card Default Model Performance Validation Results

1. The Production Model

- This section should be the same as the corresponding section in the Performance Monitoring Plan.

2. Model Development Performance

- This section should be the same as the corresponding section in the Performance Monitoring Plan.

3. Performance Monitoring Plan

- This section should be the same as the corresponding section in the Performance Monitoring Plan.

4. Performance Monitoring Results

- Produce a lift chart (in table format) and compute the Kolmogorov-Smirnov (KS) statistic for the validation data set. Use twenty groups to produce your lift chart. These groups are called semi-deciles or half-deciles.
- Compare the resulting KS statistic to the RAG (Red-Amber-Green) status. Report your recommendation to model governance.