

Class: DSE I2100
Professor: Michael Grossberg
Wayne Lam

Final Project Self-Grades

- I. Machine Learning Question: **5**. We selected a real-world machine learning problem. Explained what and why.
- II. Data Cleaning/Checking/Data Exploration: **3**. Not enough depth 1 and 2 table EDA. Could have gone a step further by comparing column correlation to loan defaults. Not enough visualizations detailing data at the table or aggregate level.
- III. Transformation, Feature Selection, and Modeling: **4**. Fairly exhaustive with transformations and checking results on model performance.
- IV. Metrics, Validation and Evaluation: **4**. Did not have ROC plotted, just reported AUC-ROC scores.
- V. Visualization: **2**. Just not a lot that really detailed the data.
- VI. Code: **3**. The final report notebook can reproduce the work from top to bottom. Needed to organize notebooks better and transfer experimental notebooks more into finalized notebooks that are easier to share results and ideas.
- VII. Presentation: **3**. A bit clumsy at times. Some results probably did not need to be shared in detail.
- VIII. Report: **3**. Too much code. All functions should have been moved to a utils file. Repeated transformations could have also been moved to functions in a utils file to reduce the amount of code in the report.