**ACCY 577 Rubric for Final Project**

The final project is worth 280 points, including the teammate evaluation.

|  |  |  |  |
| --- | --- | --- | --- |
| **Criteria** | **Full Credit** | **Half Credit** | **No Credit** |
| Group Report – Data Understanding and Feature Selection  Weight: 60 points | * Provides convincing evidence to support why features were selects, and why at least some features were not selected. * Shows proficiency in using the techniques employed (e.g., descriptive stats, groupby, pivot tables, visualization, RFE, etc.). * Considers numeric and categorical features. * All code cells run without errors. * Nicely formatted use of markdown cells to summarize the code, and demonstrate domain knowledge. | * Your work shows one or two of the following deficiencies:   + Shows basic understanding of feature exploration, but does not appear to understand what the functions are doing, or why the techniques are being used.   + Some of the code cells have errors.   + Poorly formatted and organized use of markdown cells to summarize the code.   + Shows little domain knowledge about the Lending Club business. | * Your work lacks or barely includes 3 or more of the required elements. For example:   + Cursory exploration of the data and feature selection.   + Several code cells run with errors.   + Failure to include markdown cells that explain what the code cells are doing. |
| Group Report – Modeling  Weight: 40 points | * Demonstrates proficiency in preprocessing the data (i.e., separates the data into label and data, encodes categorical features, manages missing values, splits the data using random\_state). * Demonstrates proficiency in selecting and applying at least two classifiers, one of which must be logistic regression. * Correctly optimizes the models using cross-validation. * Includes a classification report and confusion matrix. * Compares models with ROC plot. * All code cells run without errors. * Using markdown, provides an accurate, complete, nicely-formatted summary of the modeling process. * Clearly indicates which model is best and why. | * Your work has one or two of the following deficiencies:   + Incomplete or inaccurate attempt to perform the preprocessing tasks.   + Uses only one classifier, but does so correctly.   + Fails to include confusion matrix, classification report, or ROC plot.   + Errors in some of the code cells.   + Incomplete, inaccurate, or poorly formatted summary of the modeling process. | * Your work lacks or barely includes 3 or more of the required elements. For example:   + Major errors in the processing tasks.   + Fails to correctly use a classifier.   + Fails to include a confusion matrix, classification report, or ROC plot.   + Little to no summary of the modeling process. |
| Group Report – Construct Loan Portfolio  Weight: 40 points | * Correctly uses the model identified earlier to predict the loans in the test set that are likely to be paid off. * Compares the actual paid-off rate to the paid-off rate of the prediction by the model. Identifies if it’s better than the overall paid off rate of 85.2%. * Calculates the annual return and the risk of the portfolio. * Adjusts the class\_weight or other hyperparameters and evaluates the results. * Discusses other possible ways to construct a better loan portfolio. * All code cells run without errors. * Using markdown, provides an accurate, complete, nicely-formatted summary of the loan portfolio construction process. | * Your work shows one or more of the following deficiencies:   + Does not use the model identified in the modeling section to make predictions and/or makes predictions using the training data.   + Creates a portfolio, but does not evaluate the quality of the predictions using all of the methods indicated.   + Does not correctly calculate the risk of the loan portfolio.   + Inaccurate discussion of other possible ways to construct a better loan portfolio.   + Some code cells run with errors.   + Inaccurate, incomplete and/or poorly formatted summary of the loan construction process. | * Your work lacks or barely includes 3 or more of the required elements. For example:   + Fails to make predictions and create a new loan portfolio and assess the risk of the portfolio by adjusting the hyperparameters.   + There is barely a discussion of other possible ways to construct a better loan portfolio, or any markdown cells describing the loan portfolio process. |
| Group Report – Conclusion  Weight: 20 points | * Brings together the main results in a concise, well-organized, easy-to-understand way. * Makes connections among analyses, course concepts, peer-to-peer lending domain. * Recognizes the limitations of the findings, and presents possibilities for improvement. | * Your work shows one or more of the following deficiencies:   + The summary that goes into too much detail, and does not strive to present the main points.   + Barely makes connections among the analyses, course concepts, and peer-to-peer lending domain.   + Barely mentions limitations or ways to improve the analysis. | * Your work lacks or barely includes 3 or more of the required elements. For example:   + Barely includes, or fails to include a summary.   + The summary is completely unrelated to the analyses.   + Does not attempt to make connections with course content, and/or is unprofessional. |
| Video Presentation – Content  Weight: 80 points | * Begins with an agenda that clearly states the main elements of the presentation. * Discusses each of the four main tasks of the report: Data understanding and feature selection, modeling, constructing a loan portfolio, and conclusion. * Uses thoughtful transitions to connect each section and speaker in the presentation. * Focuses on key concepts and thoughtfully implements charts and graphics to highlight those concepts. * Ends with a conclusion that brings everything together and highlights the main points. | * Your presentation shows one or more of the following deficiencies:   + Lacks order by not including an agenda or conclusion.   + Does not include a discussion about one of the four main tasks.   + Does not flow well between speakers or points.   + Spends most of the time reviewing code.   + Is not connected to or supported by your work in the Jupyter notebook. | * Your presentation lacks or barely includes 3 or more of the required elements. For example:   + Fails to include an agenda and/or a conclusion.   + Does not discuss one or more of the four main tasks.   + Shares opinions that are not supported by the analyses in the Jupyter notebook.   + Spends most of the time reviewing code. |
| Video Presentation – Delivery and Use of Media  Weight: 20 points | * All group members show their faces and are actively engaged in presenting a roughly equal amount of the content. * All team members show their enthusiasm about the presentation. * The discussion is well thought out, and does not ramble. (A script may be helpful.) * The presentation of images, text, and charts are professional and have an appropriate resolution and size so that they’re easy to view. * The video presentation stays within the 8 to 12 minute limit. | * Your presentation shows one or more of the following deficiencies:   + Not all team members have a substantial role in presenting.   + There is a rambling discussion that seems to be put together on the spot.   + It’s difficult to view the images, text, and/or charts because of their size or resolution.   + The video presentation is not within the 8 to 12 minute limit. | * Your presentation shows three or more of the following deficiencies:   + Not all team members have a substantial role in presenting.   + There is a rambling discussion that seems to be put together on the spot.   + It’s difficult to view the images, text, and/or charts because of their size or resolution.   + The video presentation is not within the 8 to 12 minute limit. |
| Teammate Evaluation  Weight: 20 points | Completion of the evaluation before the deadline | Completion of the evaluation *after* the deadline | Failure to complete the evaluation |