

DS 4400: Machine Learning and Data Mining I

Spring 2022

Project Grading Rubric

- Presentation recording – 20 points
- Exploratory data analysis – 18 points
 - Info about the dataset, features, and labels
 - Discuss feature representation and selection
 - Include graphs on selective feature distributions or other visualizations
- Machine learning models – 30 points
 - Train and test at least 4 models
 - Use correct methodology (e.g., cross-validation, random splitting of training and testing set)
- Metrics – 10 points
 - Report several metrics to evaluate and compare models
 - Generate ROC curves or include confusion matrix
- Interpretation of results – 15 points
 - Why the models make errors; which features are most relevant; why is it a challenging task
- References – 3 points
 - List related literature you consulted for the project
- Peer Review – 4 points
 - Review 2 other projects and add comments to Canvas