CS200B: Introduction to Machine Learning

Project Rubric

Your code workbook will be graded out of a total of 10 points. You will receive one score for each section below, and the final score is the sum of the section scores.

Colab Notebook - Explanation of the data set.

Points	Criterion
0	Notebook contains explanation of some aspects of working with the data set, but is not clear or thorough enough to be understandable to someone unfamiliar with the project.
1	Notebook contains clear and thorough explanation of the data, including the goals in working with the data.

Colab Notebook - Exploratory data analysis.

Points	Criterion
0	Notebook does not contain any exploratory data analysis.
1	Notebook is missing one or more of the following: computation of summary statistics, principal component analysis, data visualization, or discussion of the results of these computations.
2	Notebook contains all of the following: computation of summary statistics, principal component analysis, and data visualization, and discussion of the results of these computations. However, the discussion of these results is either unclear, incorrect, or insufficient.
3	Notebook contains all of the following: computation of summary statistics, principal component analysis, and data visualization, and discussion of the results of these computations. The discussion of these results is clear, correct, and thorough.

Colab Notebook - Training models using machine learning algorithms.

Points	Criterion
0	Notebook is missing two or more of the following: a k-nearest neighbor classifier, a decision tree classifier, a random forest classifier, a support vector machine classifier.
1	Notebook includes at least three of the following: a k-nearest neighbor classifier, a decision tree classifier, a random forest classifier, a support vector machine classifier.
2	Notebook contains all of the following: a k-nearest neighbor classifier, a decision tree classifier, a random forest classifier, a support vector machine classifier. There is exploration of parameter tuning for multiple parameters and algorithms.
3	Notebook contains all of the following: a k-nearest neighbor classifier, a decision tree classifier, a random forest classifier, a support vector machine classifier. There is thorough exploration of parameter tuning for multiple parameters and algorithms, including training on different numbers of components resulting from PCA.

Colab Notebook - Analysis of Results.

Points	Criterion
0	Notebook does not contain any analysis or discussion of the results of the machine learning algorithms.
1	Notebook contains computations of the accuracy of the results.
2	Notebook contains computations of accuracy of the results for all algorithms across multiple training-testing splits, and includes discussion comparing the results.
3	Notebook contains thorough and complete computations of accuracy of the results for all algorithms across multiple training-testing splits, and includes thorough and complete discussion comparing the results.

(continued on reverse)

In-class Presentation.

Your presentation will be graded on the following aspects, for a total of 10 points.

Points	Criterion
1	Presentation is polished and well-prepared
1	Presentation tells a compelling story
1	Thorough and clear explanation of the data set
1	The goals of the project are clearly defined
1	Visualizations used make sense
1	Visualizations used are thoughtful and tell a story
1	Discussion and analysis of the results and findings
1	The presentation is given at a level easily understandable to other students in the class

In-class Reviews

You will be reviewing up to three presentations during our final, for a total of 12 points. Each review will be scored on the following aspects, totaling 4 points each:

Points	Criterion
1	Questions from the peer review have been answered
1	Some of the answers given are thoughtful and applicable
1	All of the answers given are thoughtful and applicable
1	The advice given for the presenter is insightful