In this assignment, I developed and worked on improving my implementation of the perceptron algorithm from homework 1. The differences between my model and the textbook version are that my model initializes weights as a zero-vector based on the number of features in the dataset, which initializes weights with a bias term. The textbook version also incorporates data shuffling to prevent cyclic behavior during training and includes a random seed. During the training process, my implementation updates weights based on the difference between predicted and actual outputs, the textbook model tracks misclassifications per epoch and updates both the bias and feature weights separately. Additionally, my model predicts outputs as 0 or 1, the textbook version uses -1 or 1.

To create synthetic data for testing the algorithms, I generated datasets based on baseball pitches, using pitch speed and spin rate as features to classify fastballs from off-speed pitches. I manipulated the labels to create the non-linearly separable version of the data. Which showed the limitations of the Perceptron when dealing with non-linearly separable data, as evidenced by its struggle to converge and perform well in this scenario dropping from 100% accuracy to 50%.

While implementing the Adaline method, I encountered several challenges. Initially, I struggled with the textbook implementation but eventually used functions from the sklearn such as train\_test\_split to ease the process. For the Titanic dataset using the Adaline model, I recorded a training accuracy of 62.92% and a test accuracy of 58.58%. Despite this my efforts to analyze the weights and identify the most predictive features, I was unable to extract meaningful outputs, with all my weight values returning as NaN.

To compare my models against a baseline, I implemented a simple baseline model that achieved a training accuracy of 60.35% and a test accuracy of 56.72%. This comparison demonstrated that both the Perceptron and Adaline models outperformed the baseline, reinforcing the notion that they were learning from the data rather than making random guesses.