In this class, we learned about the different metrics used for model evaluation and their upsides/downsides. In linear regression, we generally consider mean squared error (MSE), mean absolute error (MeaAE), and median absolute error (MedAE). MSE is the standard metric that finds the average "distance" between the prediction and the data. Because it is squared, it augments the penalization for far outliers (useful for converging the gradient descent faster in multidimensional case when doing matrix inversion is not that easy). Disadvantage: when using it as a loss function, the linear regression becomes sensitive to large outliers because we would have to deal with their "squaredness." MeaAE is less sensitive to outliers and is able to find the average residual. However, when the distribution of residuals is skewed from (-x to +x), MeaAE alone will not be able to characterize the behavior of the regression. Here comes MedAE which shows where the peak of the distribution of residuals is. A disadvantage to this metric is that when the distribution is multimodal, it might be pointless.

For classification tasks (e.g., kNN, logistic regression, etc.), we used metrics such as accuracy, precision, recall, and ROC. While accuracy is widely used, it is misleading when out dataset is skewed since we will have a large percentage of FP or FN that would be small in terms of absolute count, making the accuracy higher than it "actually is." Here come precision (the ability to find data points only of class 1) and recall (the ability to find all data points of class 1). Usually, these measures are then combined into F-1 score with which we can find the best ratio for precision and recall. Another metric, ROC, helps us determine the amount of overlap between two distributions of data where the overlaps are false positive or false negatives. We use it to determine how well we're able to differentiate 2 classes.