

- a. It is important that we have multiple aggregation options because it allows us to do both classification and regression. The 'mode aggregator is suited for classification, while the 'mean' aggregator is suited for regression.
- b. A third aggregator could be 'median' – this would simply take the median value of the nearest neighbor labels and return that for a given element of testing data. This would generally be suited for regression. This could be preferable to 'mean' in cases where there are exist some outliers in our data that we would prefer to ignore. One drawback is that if such outliers are actually meaningful, it could fail to take them into account when making predictions.