Up to this point, we have focused on methods that fit parametric functions---like polynomials and hyperplanes---to the entire dataset. In this module, we instead turn our attention to a class of "nonparametric" methods. These methods allow the complexity of the model to increase as more data are observed, and result in fits that adapt locally to the observations.

We start by considering the simple and intuitive example of nonparametric methods, nearest neighbor regression: The prediction for a query point is based on the outputs of the most related observations in the training set. This approach is extremely simple, but can provide excellent predictions, especially for large datasets. You will deploy algorithms to search for the nearest neighbors and form predictions based on the discovered neighbors. Building on this idea, we turn to kernel regression. Instead of forming predictions based on a small set of neighboring observations, kernel regression uses all observations in the dataset, but the impact of these observations on the predicted value is weighted by their similarity to the query point. You will analyze the theoretical performance of these methods in the limit of infinite training data, and explore the scenarios in which these methods work well versus struggle. You will also implement these techniques and observe their practical behavior.