

A Visual Introduction to Machine Learning

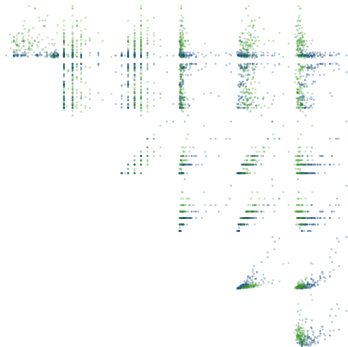
In machine learning, computers apply **statistical learning** techniques to automatically identify patterns in data. These techniques can be used to make highly accurate predictions.

Keep scrolling. Using a data set about homes, we will create a machine learning model to distinguish homes in New York from homes in San Francisco.

SCROLL

First, some intuition

Let's say you had to determine whether a home is in **San Francisco** or in **New York**. In machine learning terms, categorizing data points is a **classification** task.

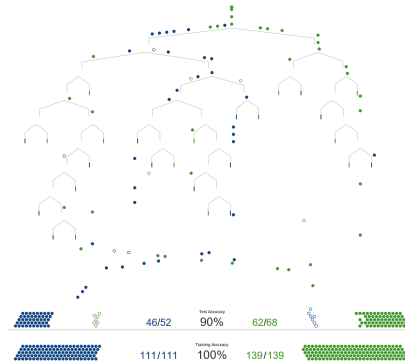


(a) scroller narrative flow

Reality check

Of course, what matters more is how the tree performs on previously-unseen data.

To **test** the tree's performance on new data, we need to apply it to data points that it has never seen before. This previously unused data is called **test data**.



(b) stepper narrative flow