Training will be based on x-train and y-train. X test is the one that is going to predict. Now this is something that has been predicted by the model. We will call this y predict and compare it with y test to check the performance of the model.

**Knn.fit:**

Fit will be for training the model (x train and y train). For unsupervised learning we will only use x train because y won’t be present.

For testing accuracy:

y-test and y-predict

For training accuracy:

x-train and y\_train

Evaluation is based on unseen data.

.score method will work on .predict and .accuracy

Generalized model is the one which has good training and testing accuracy. 0.1 (10%) difference.

Over fit mode is the one whose training accuracy is good and testing accuracy is not good. Take the example of a student who only rote learns. He can work very well on questions he’s already seen because he’s learned them but on unseen data, he cannot work.

Under fit model is the one whose training accuracy is not good and testing accuracy is also not good.

Accuracy will change because of random state too.

**True positive:**

A patient who had covid positive result, model said that he is covid positive

**True negative:**

A patient who had covid negative result, model said that he is covid negative

**False positive:**

A patient who had covid positive result, model said that he is covid negative

**False negative:**

A patient who had covid negative result, model said that he is covid positive