

**Overfitting:**

Overfitting refers to a model that models the training data too well.

Overfitting happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. This means that the noise or random fluctuations in the training data is picked up and learned as concepts by the model. The problem is that these concepts do not apply to new data and negatively impact the models ability to generalize.

Overfitting is more likely with nonparametric and nonlinear models that have more flexibility when learning a target function. As such, many nonparametric machine learning algorithms also include parameters or techniques to limit and constrain how much detail the model learns.

Example when k = 1, here our model is making too much effort to not skip even a single point, that may be outliers in way of making perfect training, but in such case it leads to consider outliers too as in our example, one orange is there blue region, it also consider it even though it’s an outlier.

**Underfitting:**

Underfitting refers to a model that can neither model the training data nor generalize to new data.

An underfit machine learning model is not a suitable model and will be obvious as it will have poor performance on the training data.

In this model is not making any effort and directly stating class of data without looking at data.]

Example when k=n, model clearly states the class of new datapoint as the majority class in training.

**A Good fit:**

It’s a balance between underfit and overfit or sweet spot, The sweet spot is the point just before the error on the test dataset starts to increase where the model has good skill on both the training dataset and the unseen test dataset.