When evaluating the model’s performance on the test data that the model has not seen before, we want to measure the difference/error between the model’s predictions and the true values. This error can be split into three terms.

The bias term is a quantity that describes how much the data is shifted from the mean value of the model. A high bias indicates that the model is not able to represent the patterns in the underlying data.

Variance, on the other hand, represents the variance of the predicted data from its corresponding mean, and it describes the sensitivity of the model to the noise in the data. Thus, high variance indicates that the model has not only learned the underlying model, but went further to even capture the noise in the data.

A model that is simple and does not capture the underlying trends in the data is called an underfitting model, and has high bias and low variance. On the other hand, a model that is too complex will likely result in overfitting, in which case it would have low bias and high variance.

Last is the irreducible error term. This represents the error that is not related to the complexity of the model and that cannot be reduced no matter how good and tuned the model is, with a given set of predictors. This error term is given by the variance of the noise in the data, and can be reduced by recognizing more independent predictors that are also related to the dependent variable.

This figure (add figure for bias variance tradeoff) shows the trends in the bias and the variance as a function of the model’s complexity.

As expected, at low degrees of complexity, the model suffers from under-fitting as a low order polynomial cannot reasonably represent a complex function with exponential and quadratic terms. This is shown by the high bias for polynomials of low order.

As the complexity increases, the bias starts decreasing indicating that the model is no longer deficient of approximating the overall shape and trends in the data. At a certain polynomial degree (add the exact degree here), the model reaches a balance between bias and variance, where the overall error is the lowest. Such a model is a model that has learned the trends in the data but has not yet entered the stage of memorizing the training data and adapting itself to the noise/idiosyncrasies in the training data.

When the polynomial degree exceeds (add the exact degree), the variance starts to increase in value indicating that the model has now entered into the realm of over-fitting. This variance will keep increasing as the complexity increases until the model has learned all the noise in the training data. This additional learning negatively impacts the model’s performance on new data and inhibits its ability to generalize to the domain of interest, which is the ultimate goal of any machine learning algorithm.