When dealing with test data, we want to measure the difference between the model’s predictions and the new data, which the model has not been trained on. 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.