Print

Model Complexity

A naive modeler might assume that the most complex model should always outperform the others, but this isn't the case. An overly complex model might be too flexible, and lead to overfitting, accommodating nuances of the random noise or the chance relationships in the sample. Overfitting leads to models that have higher variance when they are applied to a population.

On the other hand, an insufficiently complex model might not be flexible enough. This leads to underfitting. That is, systematically missing the signal, and underfitting leads to biased predictions. A model with just enough complexity, which also means just enough flexibility, gives the best generalization to new data sets. The important concept is that there isn't one perfect model. There's always a balance between overfitting and underfitting.

Close