

Bias- Variance Dilemma and No. of Features

high bias

pays little attention to data
oversimplified

high error on training set
(low r^2 , large SSE)

few features used



high variance

pays too much attention to data
(does not generalize well)
overfits

much higher error on test set
than on training set

carefully minimized SSE

