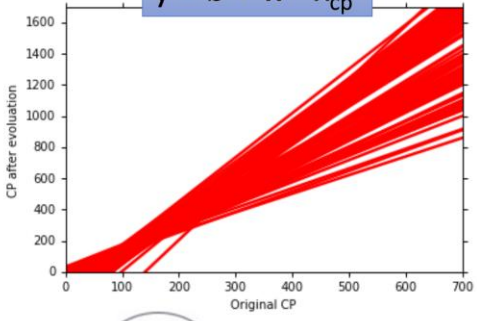
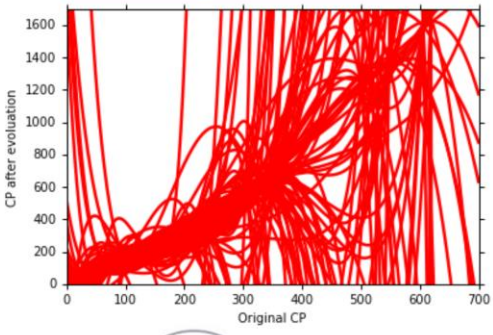
Variance

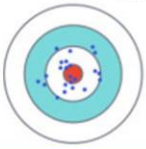
$y = b + w \cdot x_{cp}$



$y = b + w_1 \cdot x_{cp} + w_2 \cdot (x_{cp})^2 + w_3 \cdot (x_{cp})^3 + w_4 \cdot (x_{cp})^4 + w_5 \cdot (x_{cp})^5$



Small Variance



Large Variance

Simpler model is less influenced by the sampled data

Consider the extreme case f(x) = c

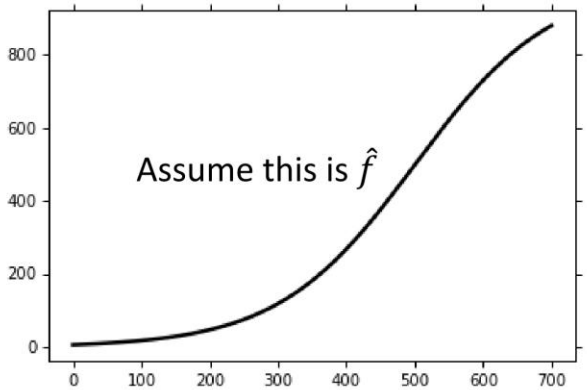
Bias

$$E[f^*] = \bar{f}$$

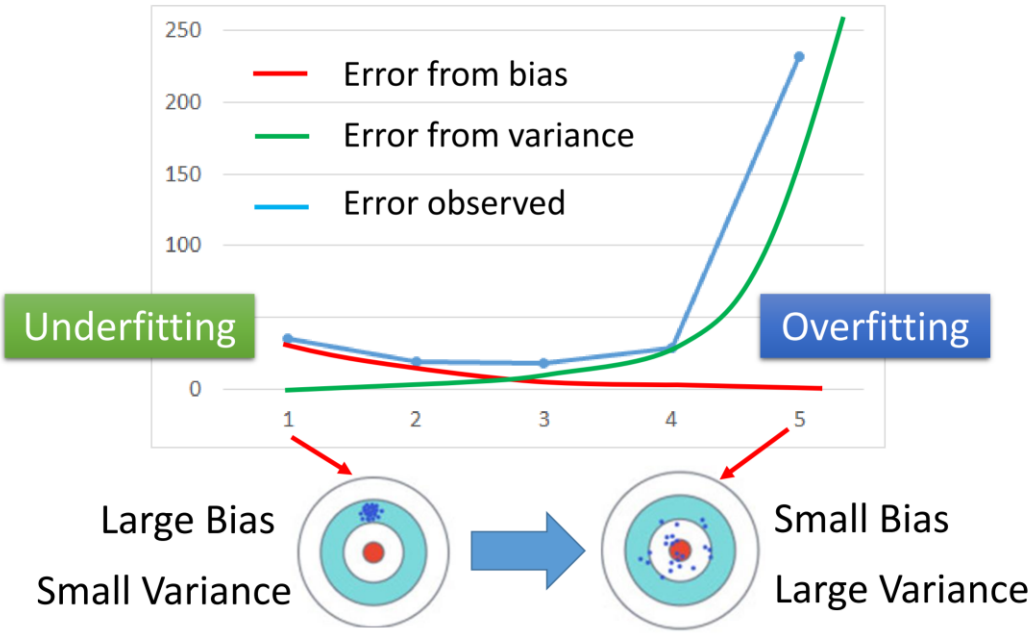
- Bias: If we average all the f^*, is it close to $\hat{f}$

Large Bias

Small Bias

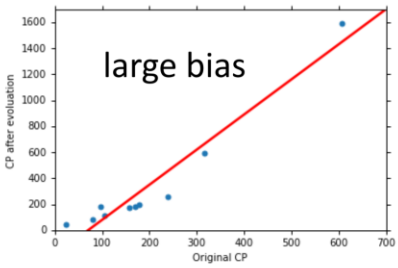


Bias v.s. Variance



What to do with large bias?

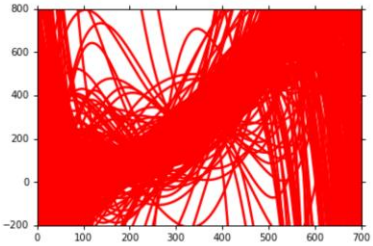
- Diagnosis:
  - If your model cannot even fit the training examples, then you have large bias **Underfitting**
  - If you can fit the training data, but large error on testing data, then you probably have large variance **Overfitting**
- For bias, redesign your model:
  - Add more features as input
  - A more complex model



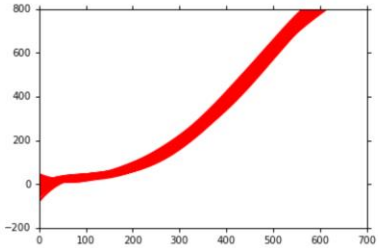
What to do with large variance?

- More data

Very effective, but not always practical

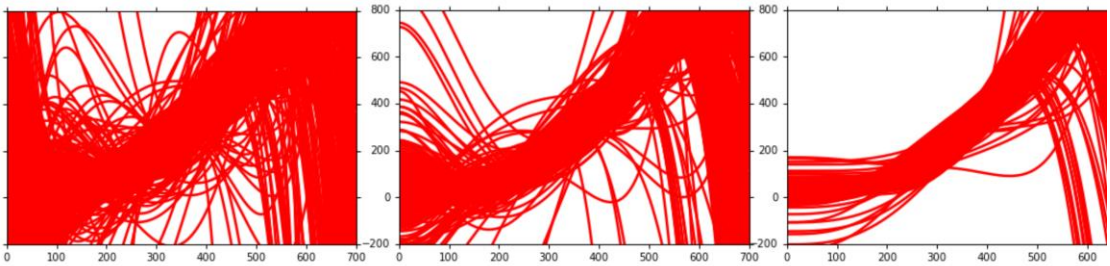


10 examples

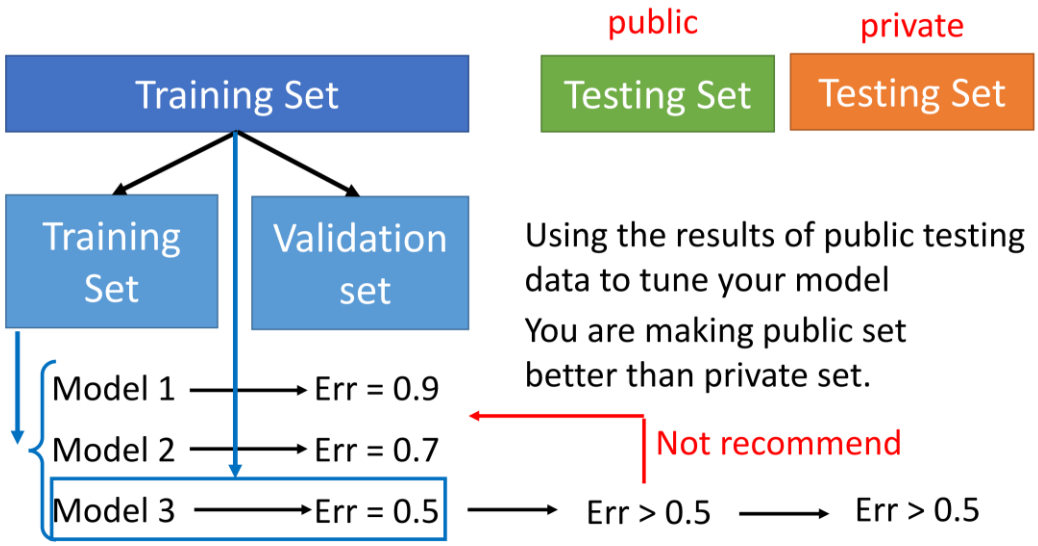


100 examples

- Regularization



Cross Validation



N-fold Cross Validation