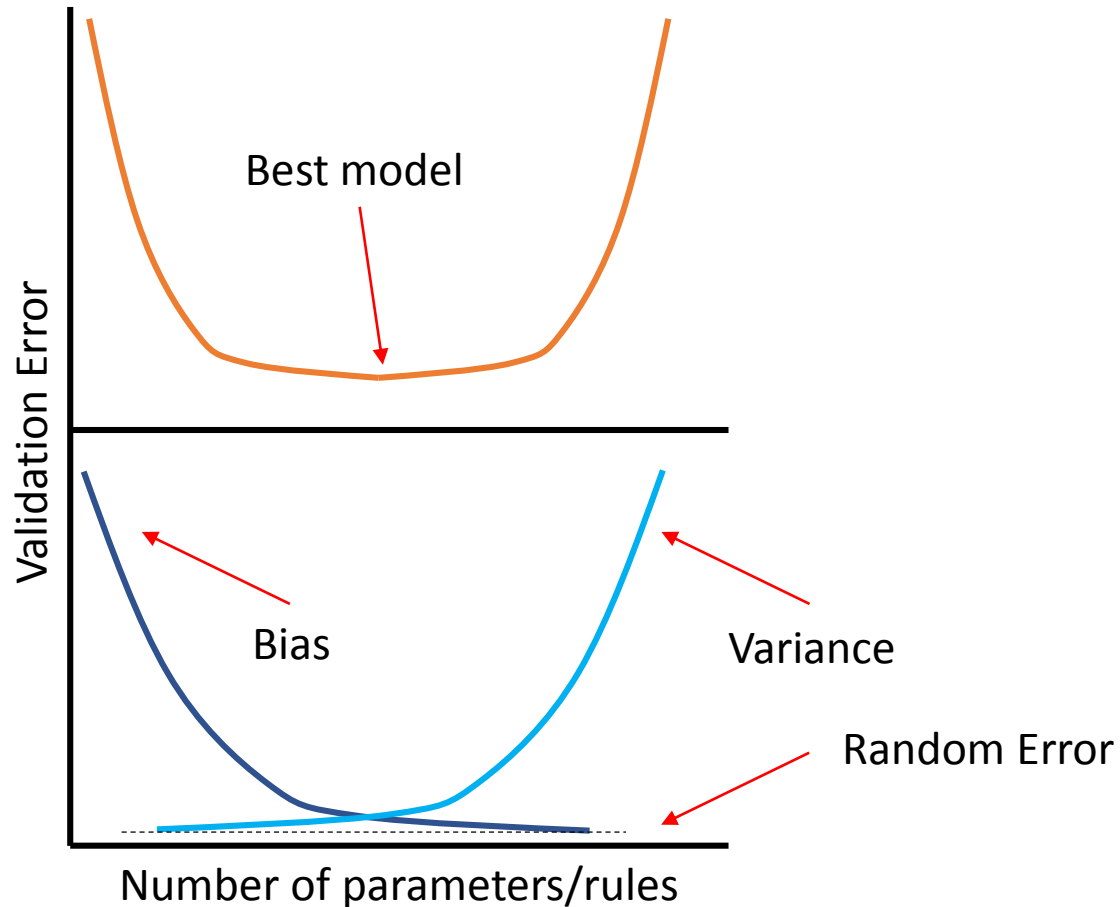


# The Bias / Variance Trade-off

Total Error = Bias + Variance + Random Error =  
 $(\hat{f}(x) - f(x))^2$



Bias =  $E[\hat{f}(x)] - f(x)$  or the error that arises from a model's inability to replicate the fundamental phenomena represented by a data set.

Variance =  $(\hat{f}(x) - E[\hat{f}(x)])^2$  or the error that arises from a model's ability to produce differing predictions from the values in a new data set.

# The Bias / Variance Trade-off: predictions



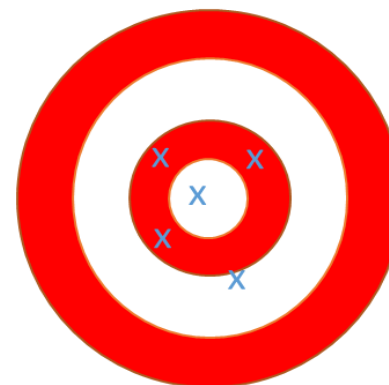
**High Bias  
Low Variance**



**High Bias  
High Variance**

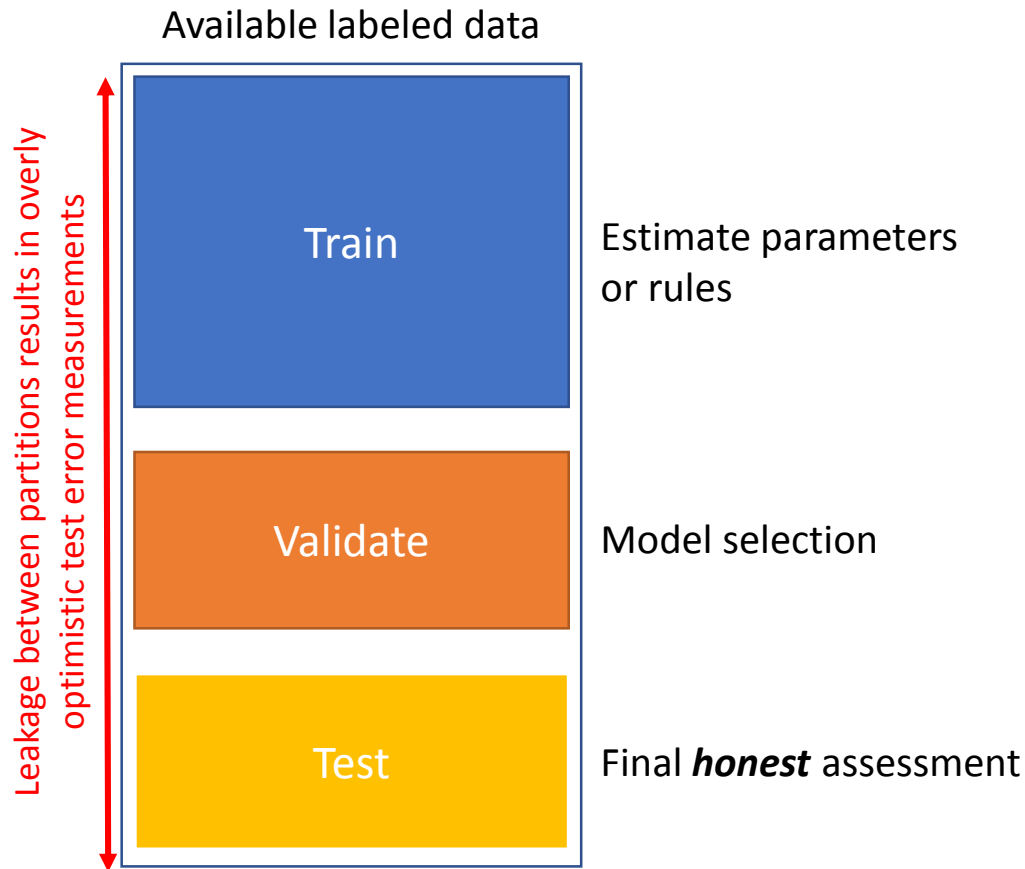


**Low Bias  
Low Variance**

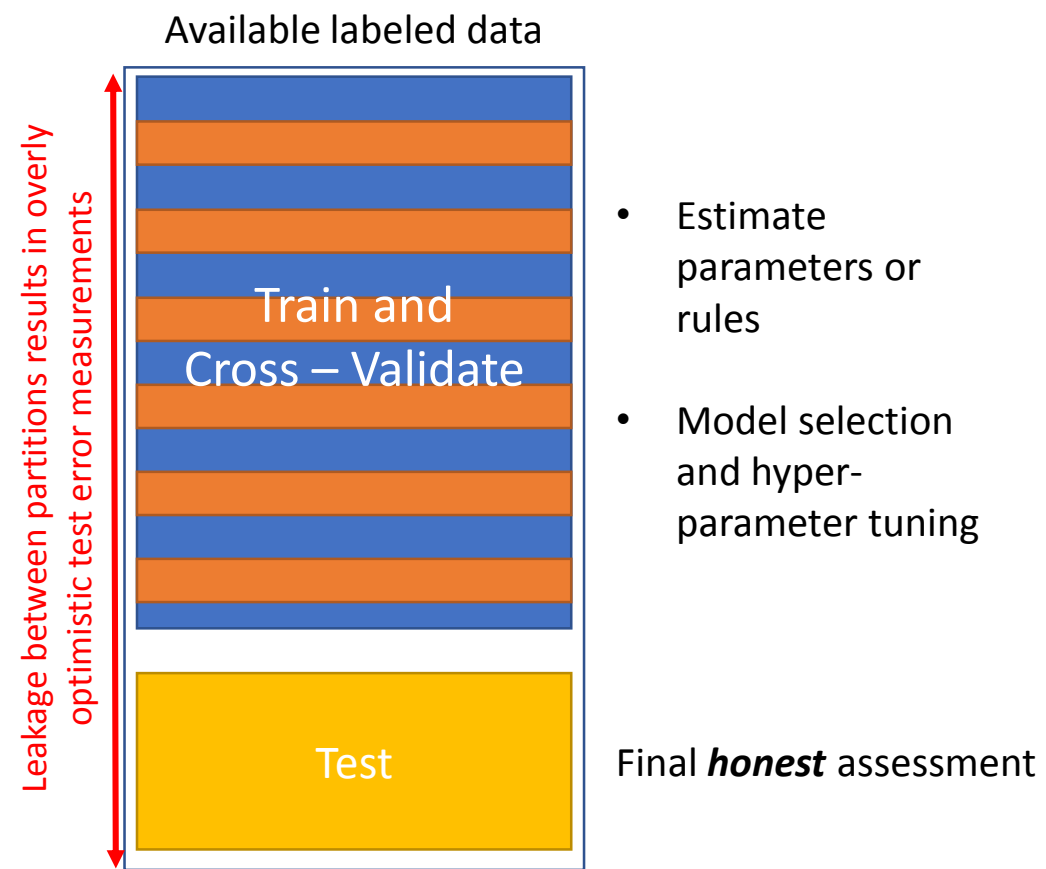


**Low Bias  
High Variance**

# Bias / Variance trade-off in-practice: Honest assessment



Best suited for big data or linear models using traditional forward, backward, or stepwise selection.



Nearly always a more generalizable approach, but computationally intensive. Best suited for complex models with many hyper-parameters and small to medium sized data.