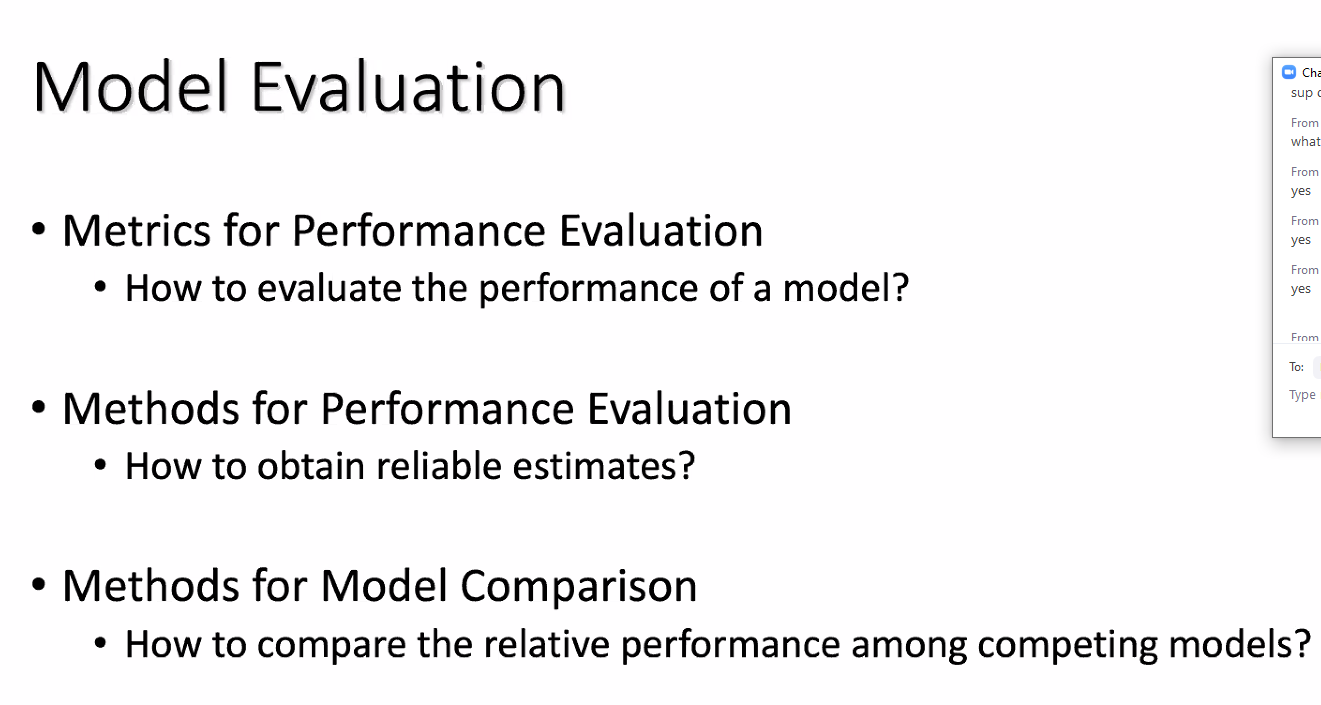
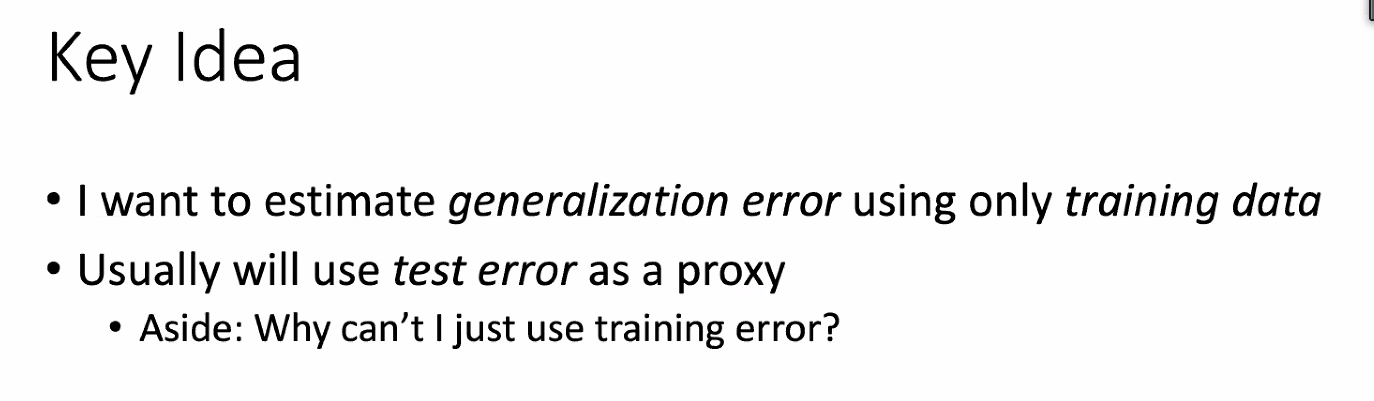
MODEL SELECTION, IMPORTANT FOR EVALUATING:



* FIRST, Methods for model comparison, how to compare the relative performance among competing models?



Use proxy for generalization error

“the test error estimated by evaluating on a held-out test set

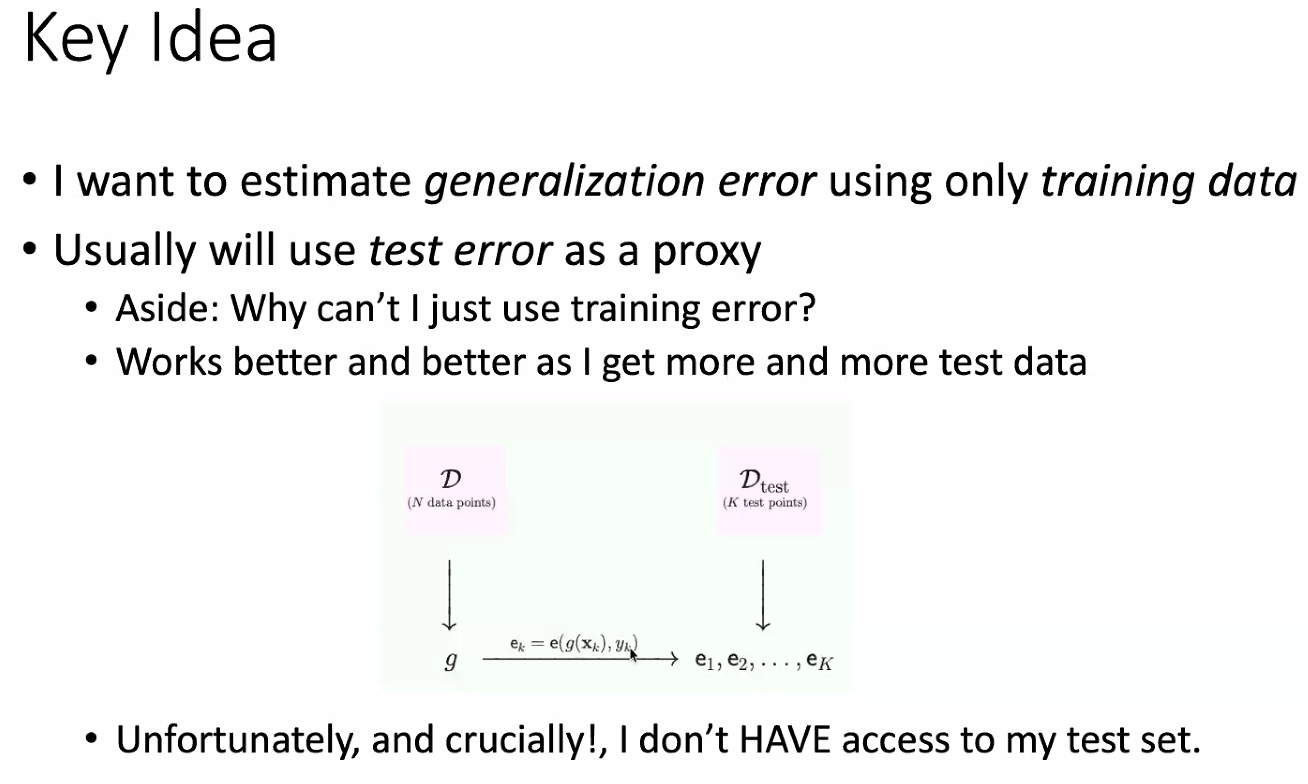
the probability that the hypothesis is not equal to the target function

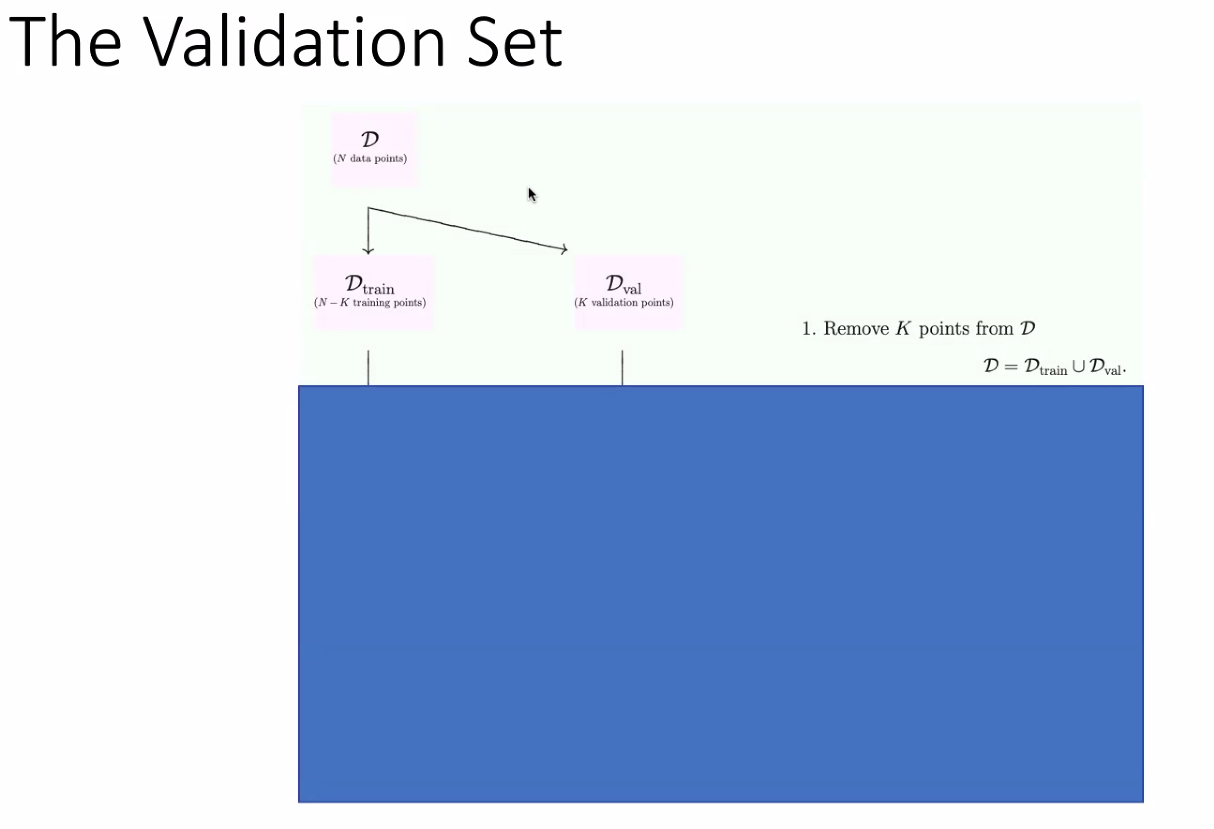
ah got it!”

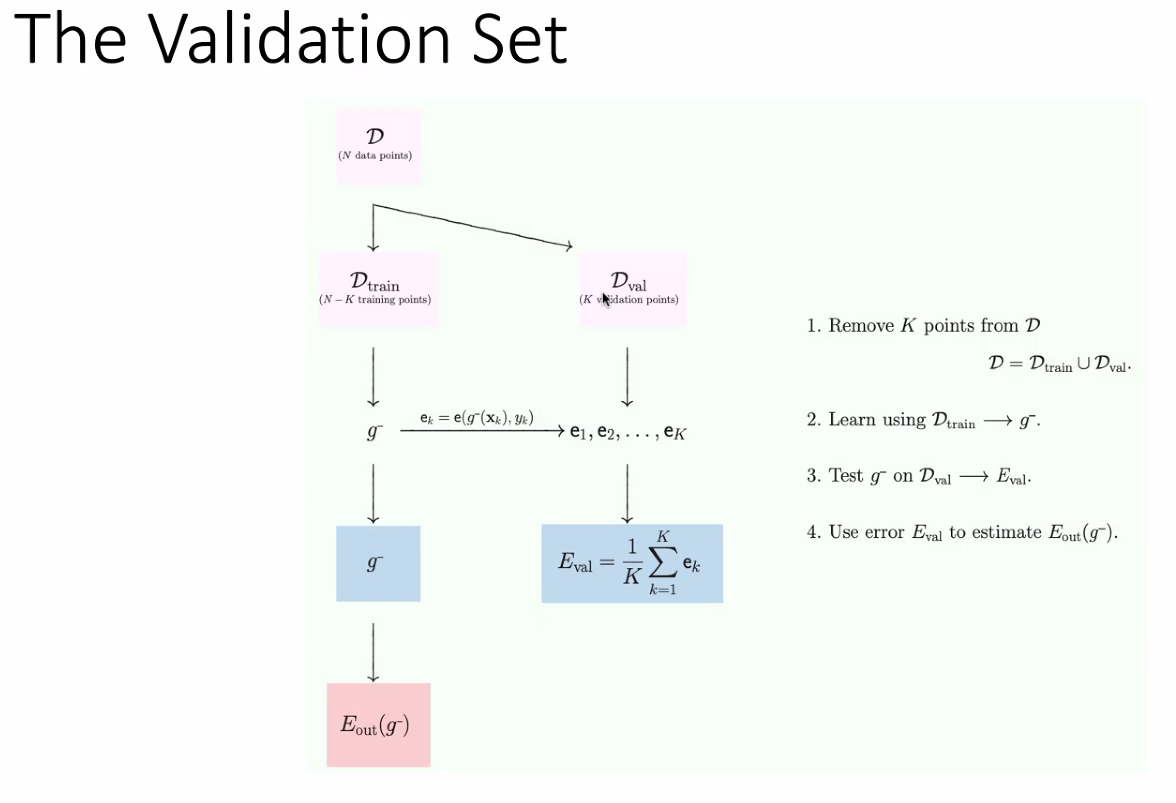
**Mini Quiz:**

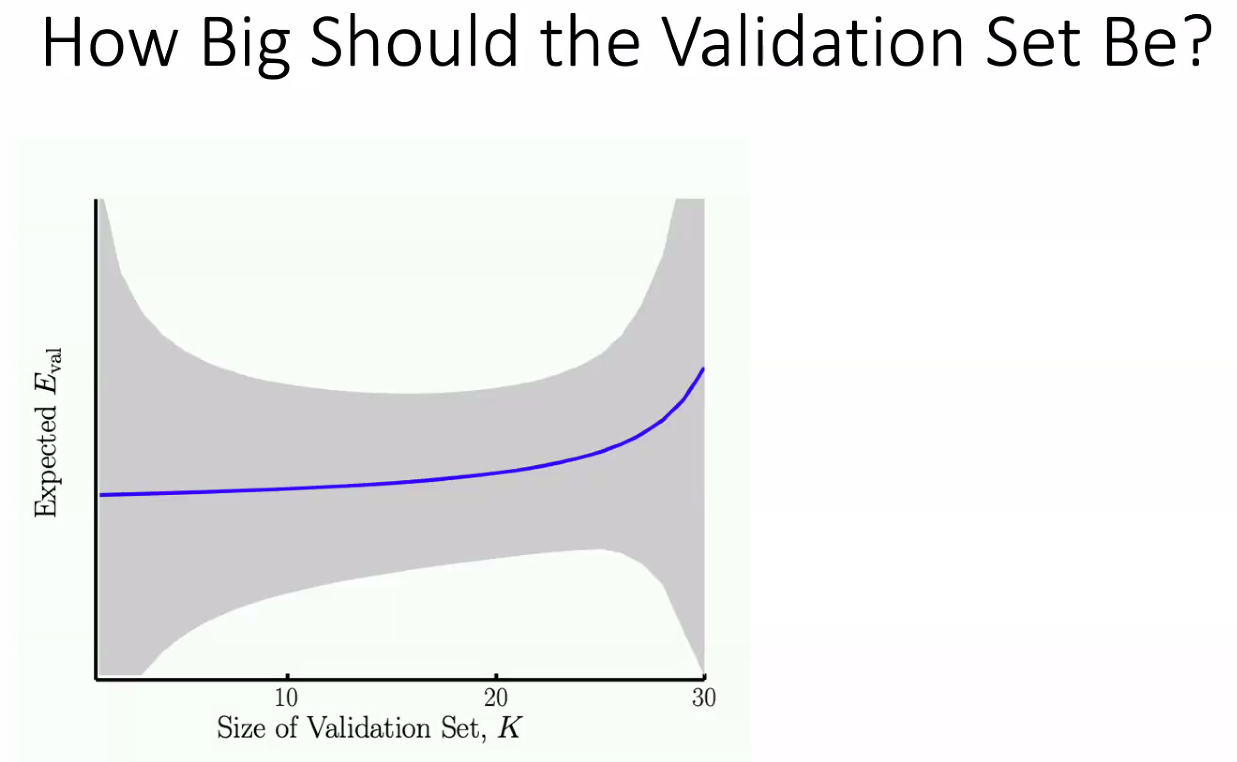
**Suppose I estimate test error using training error. Which is true?**

**Head and Tails sequence – train data underestimates the error on the test data**



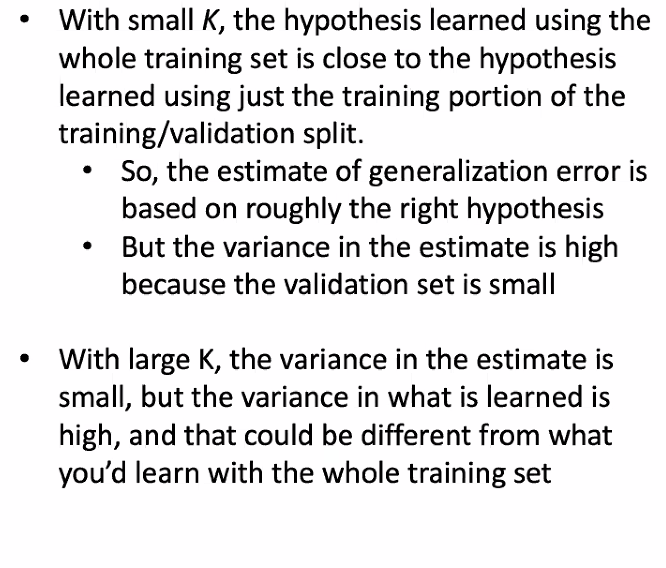




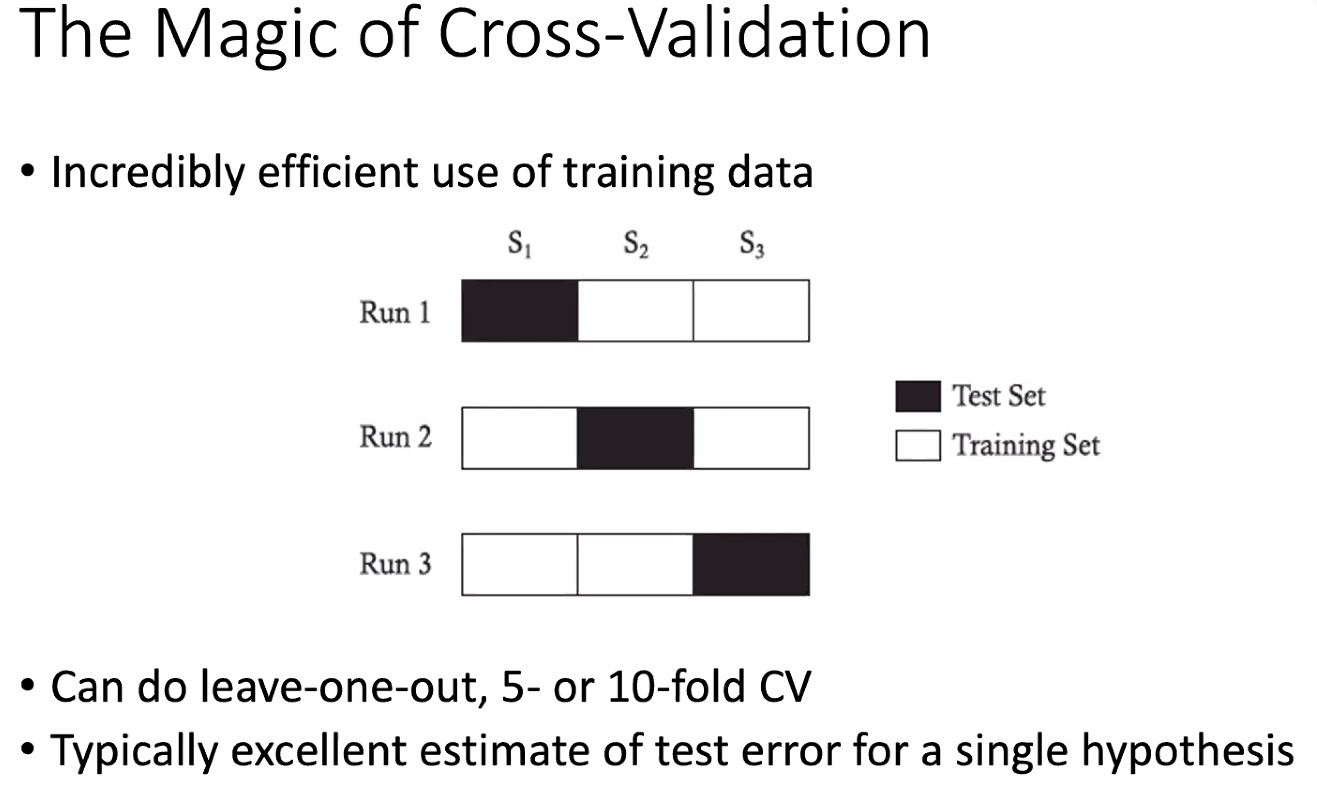


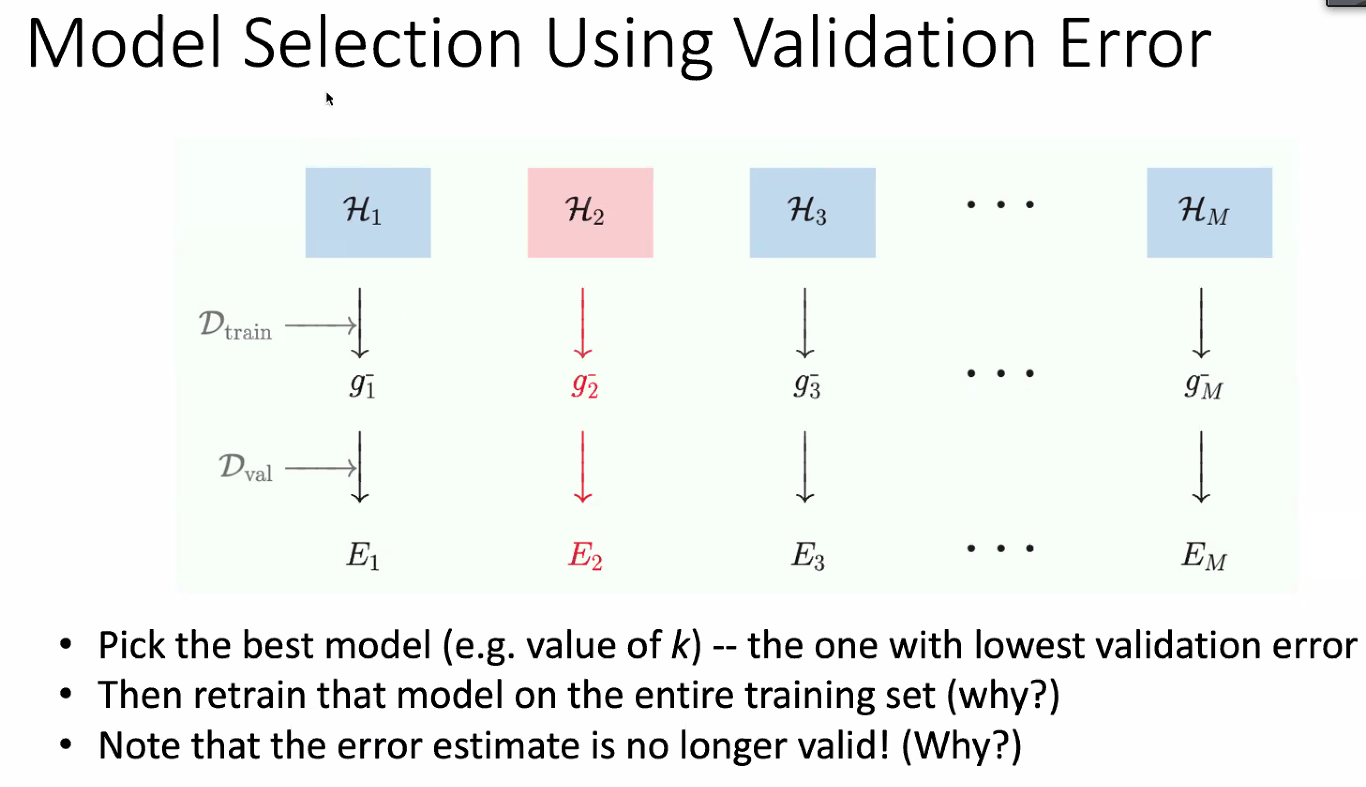
**gray part is the variance ,**

**when you have a small validation set , you have low Eval high variance, Why?** **bc the low number of data points each then have a heavier weight on the result for the validation set**

**what This means^^** 

**another Solution that’s Great:**





**Answers to those whys^^:**

1. **Make use of all of the data available to you for validation “bigger training set = more accurate results”**
2. **because you are using the validation test set for error estimation on the training data**