Model Evaluation

ML Instruction Team, Fall 2022

CE Department
Sharif University of Technology

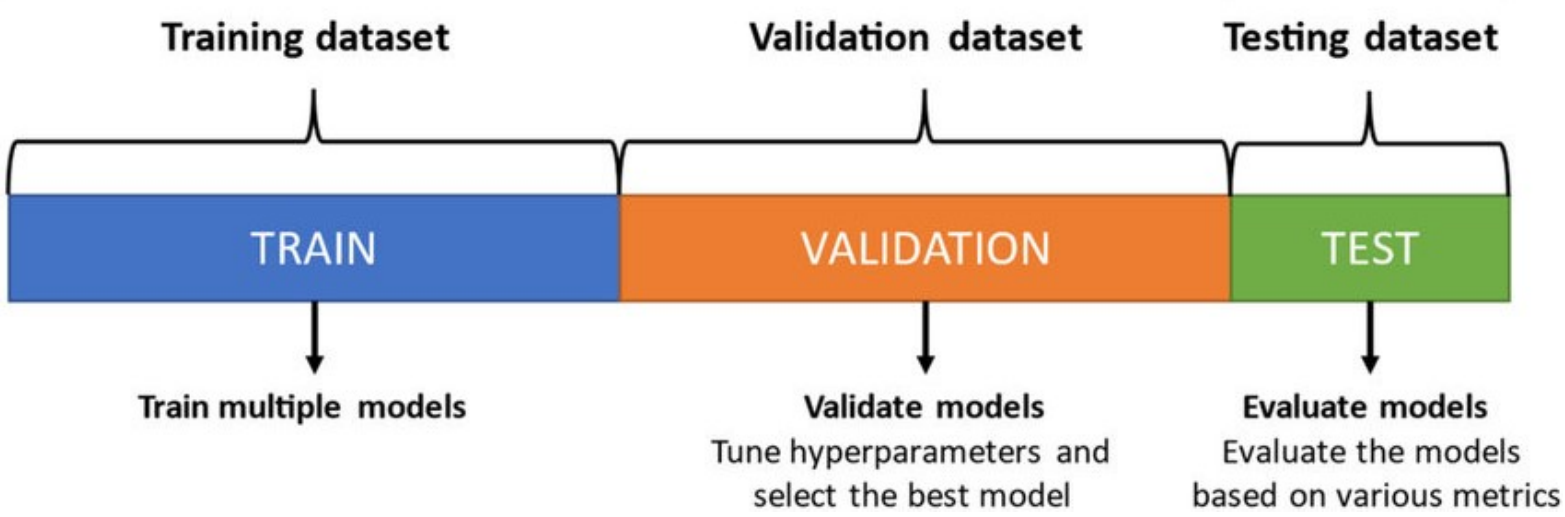
Why Evaluation?

- Estimation of the generalization error
- Increasing of the predictive performance
- Selecting best-suited ML algorithm for our problem

Why Validation?

- Training set error is an optimistically biased estimator of the generalization error
- Test set error is an unbiased estimator of the generalization error

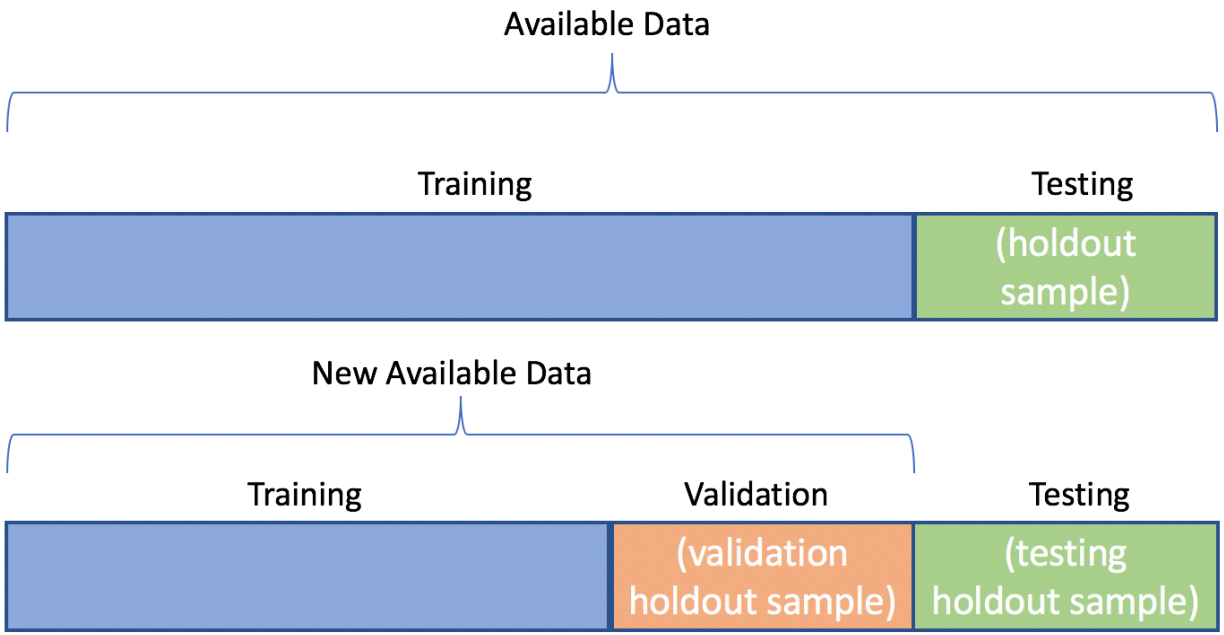
Validation


Training dataset
Validation dataset
Testing dataset
TRAIN
VALIDATION
TEST
Train multiple models
Validate models
Tune hyperparameters and select the best model
Evaluate models
Evaluate the models based on various metrics

Types of Validation

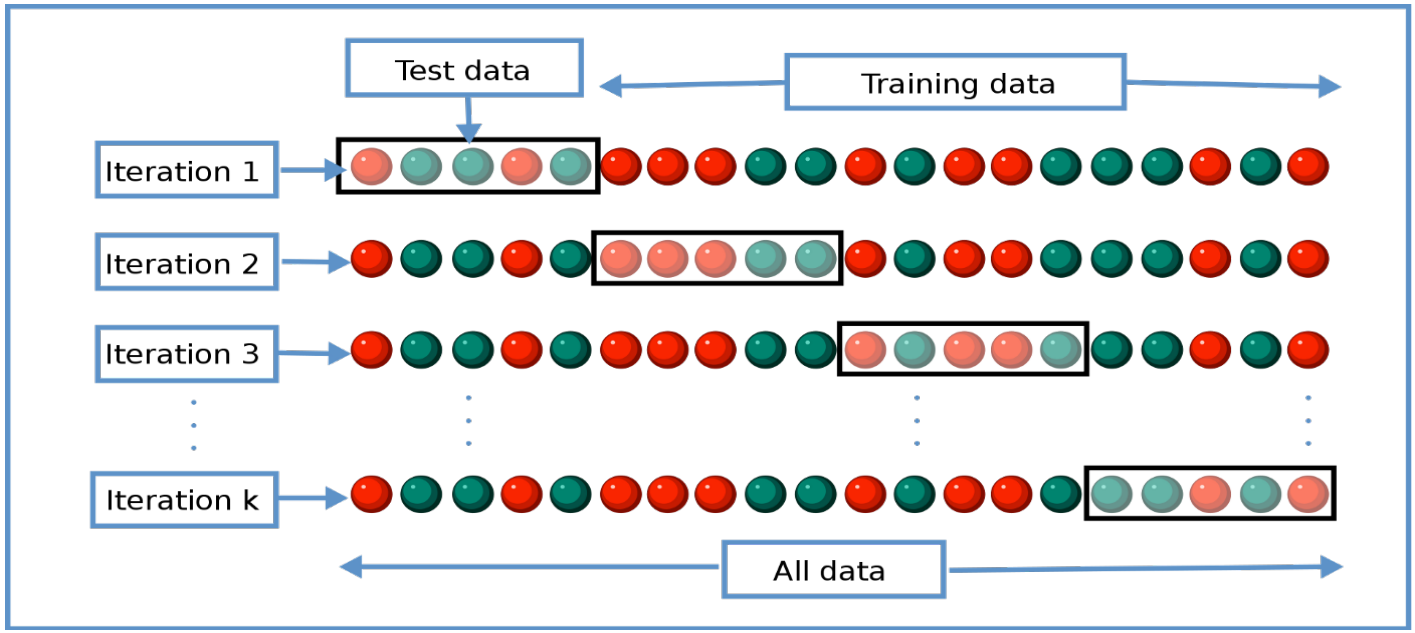
- Holdout Validation
- LOOCV (Leave One Out Cross Validation)
- K-Fold Cross Validation

Holdout Validation



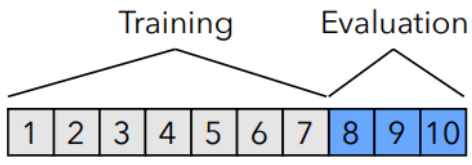
- Pros
  - Fully independent of data
  - Lower computational costs
- Cons
  - higher variance

K-Fold Cross Validation

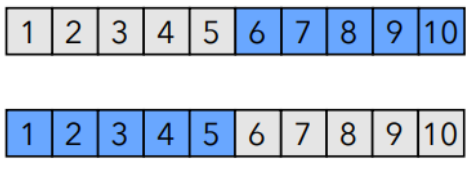


- Pros
  - Lower computational compared to LOOCV
  - Lower variance campared to Holdout
  - Reducing both Bias and Variance
- Cons
  - Higher computational costs in big data state
  - Impact model in imbalanced data state

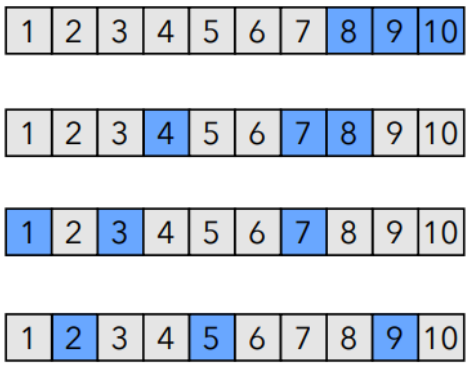
Types of Validation



Holdout Method



2-Fold Cross-Validation



Repeated Holdout

Thank You!

Any Question?