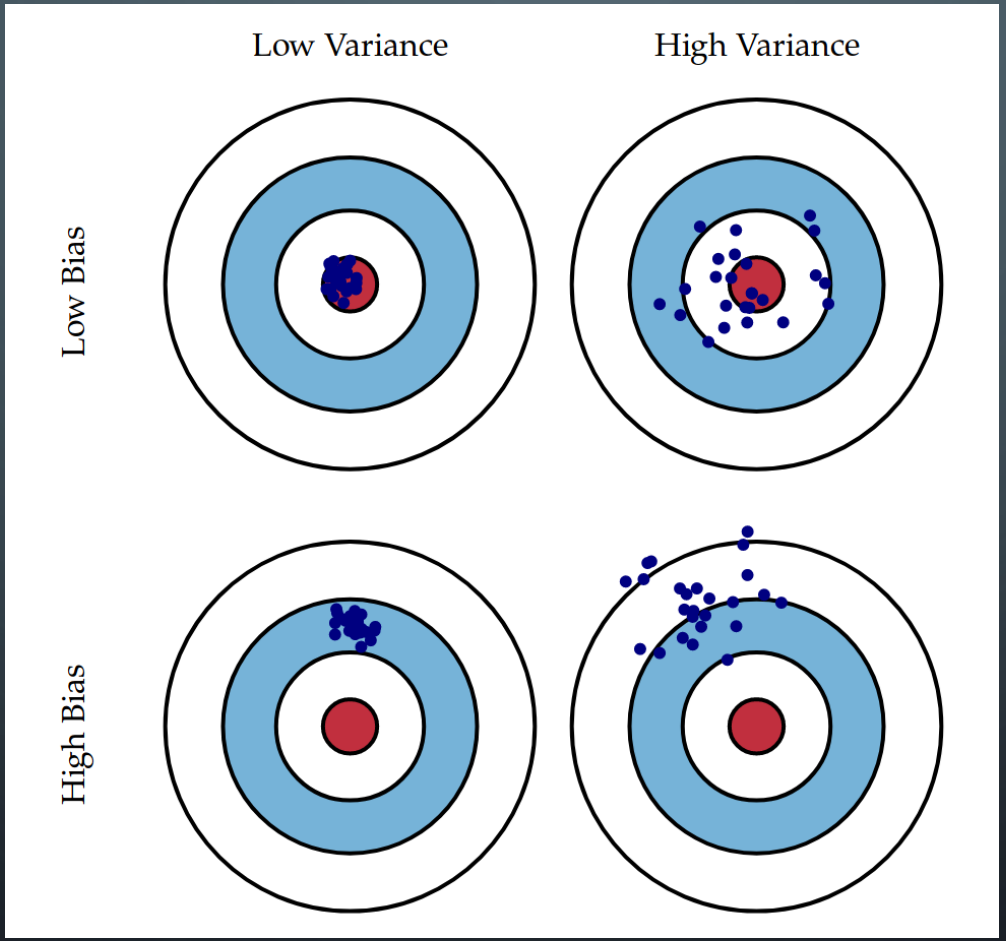
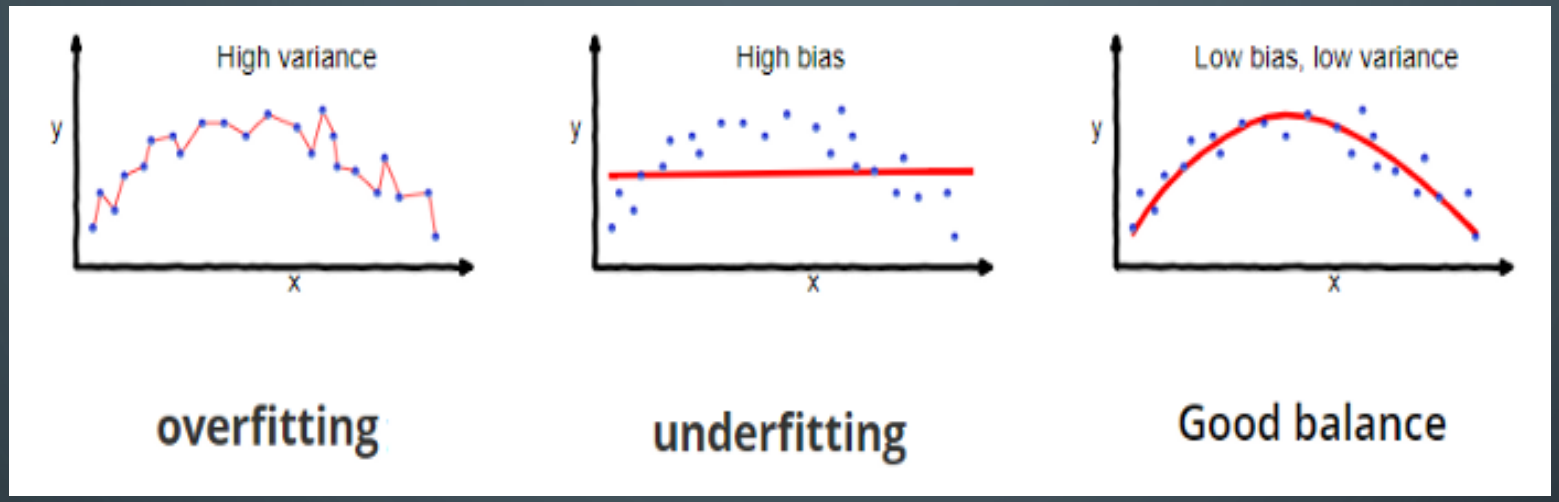
EVALUATION AND TUNING

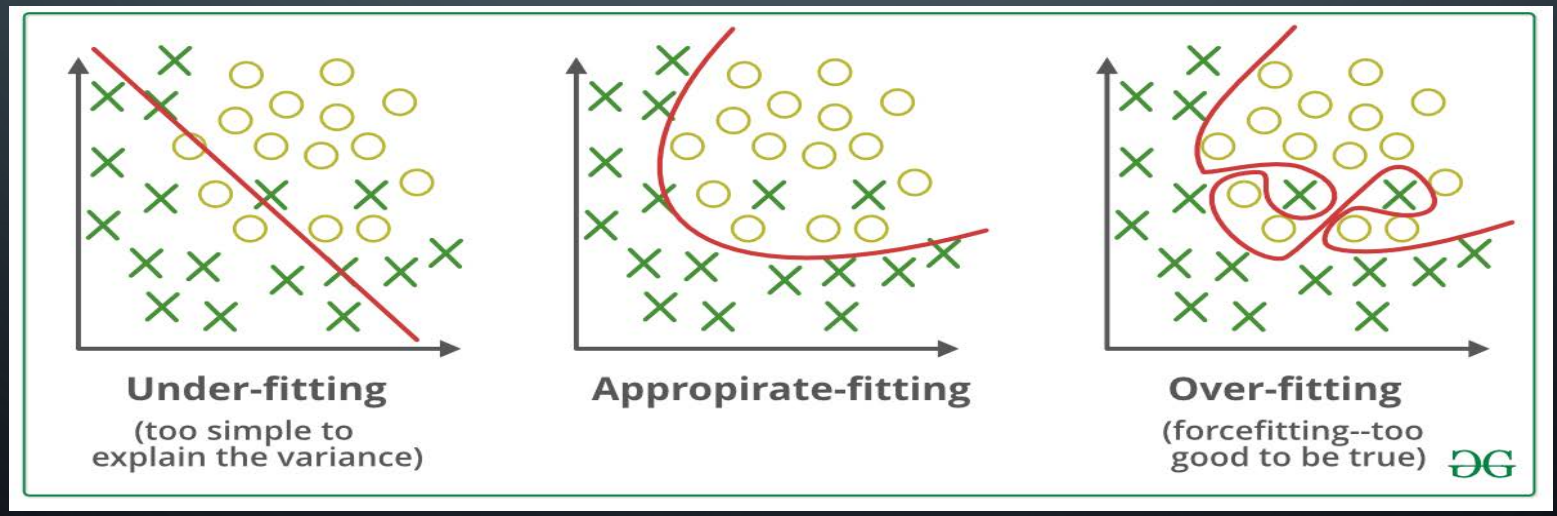
MOHAMMAD GHODDOSI

BIAS AND VARIANCE

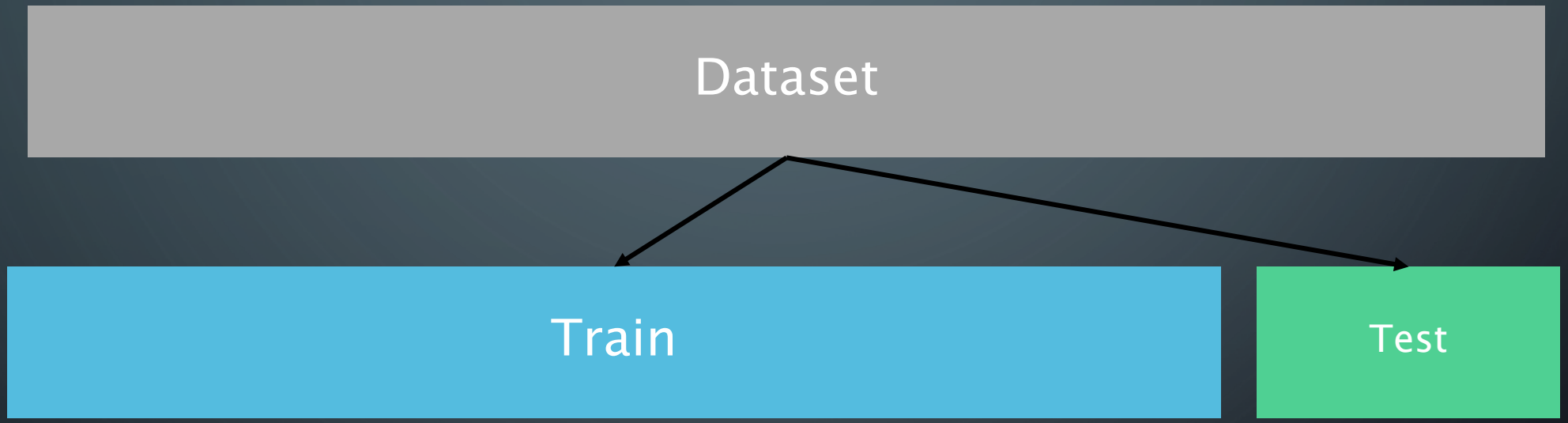


OVERFITTING AND UNDERFITTING

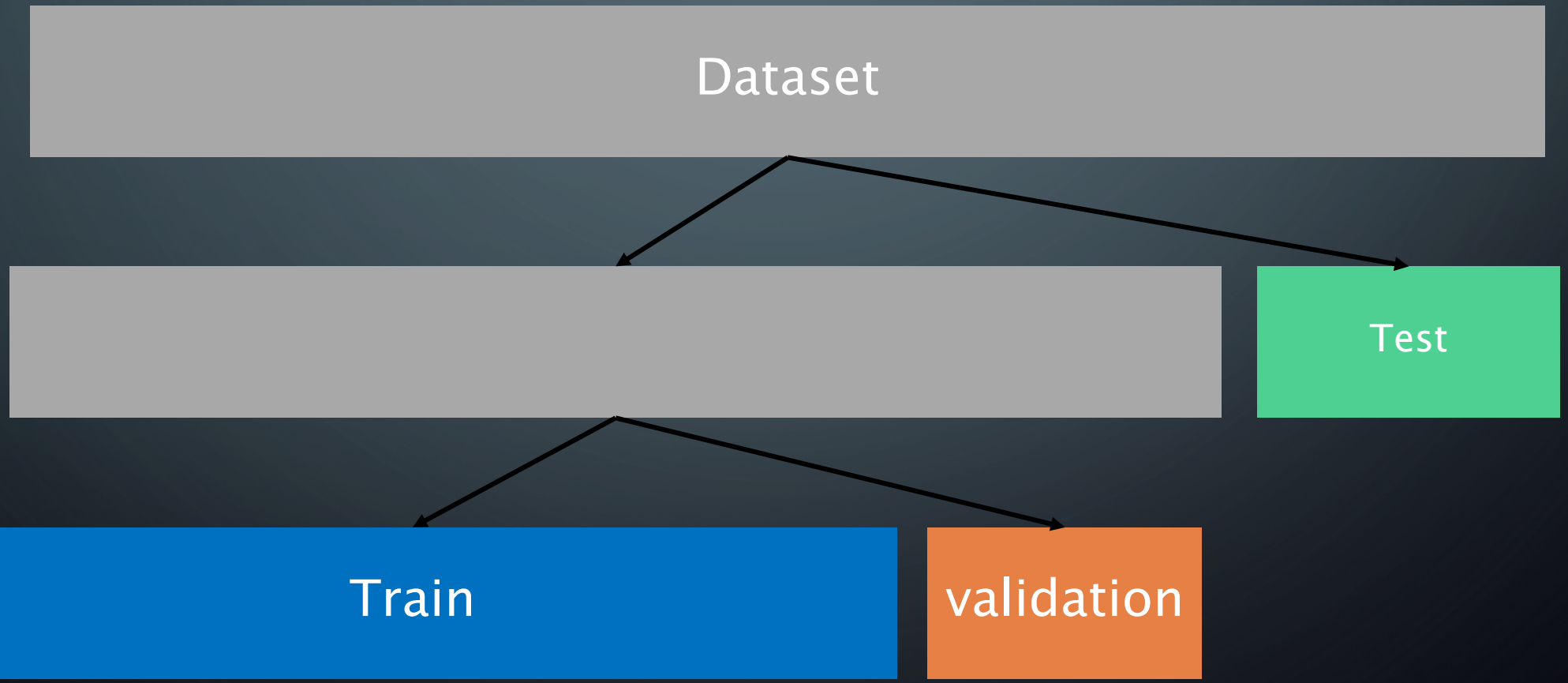




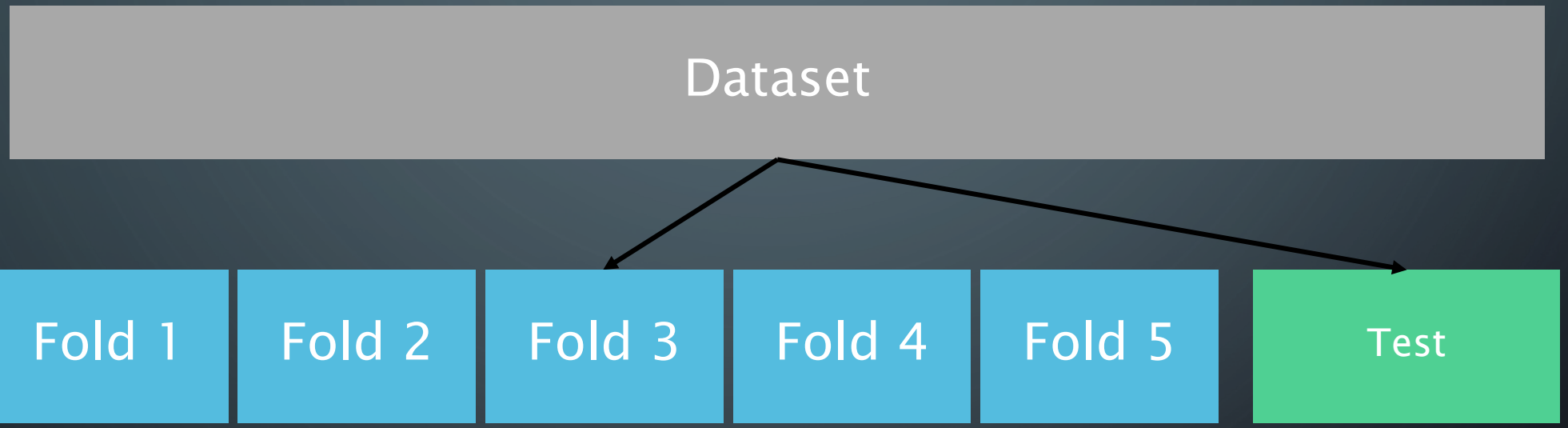
TRAIN-TEST SPLIT



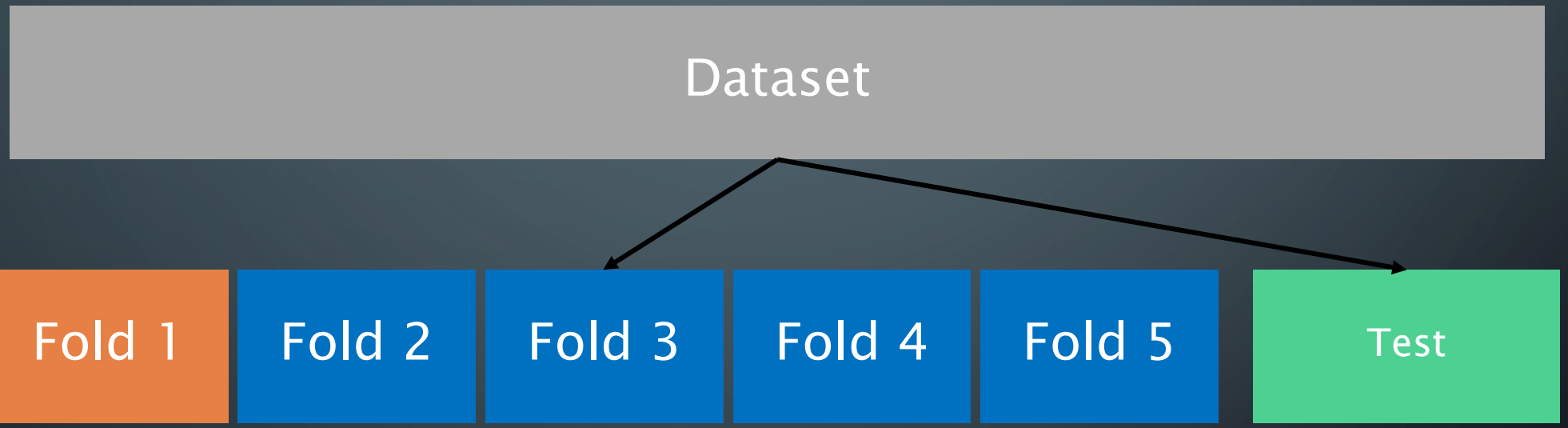
TRAIN-TEST - VALIDATION SPLIT



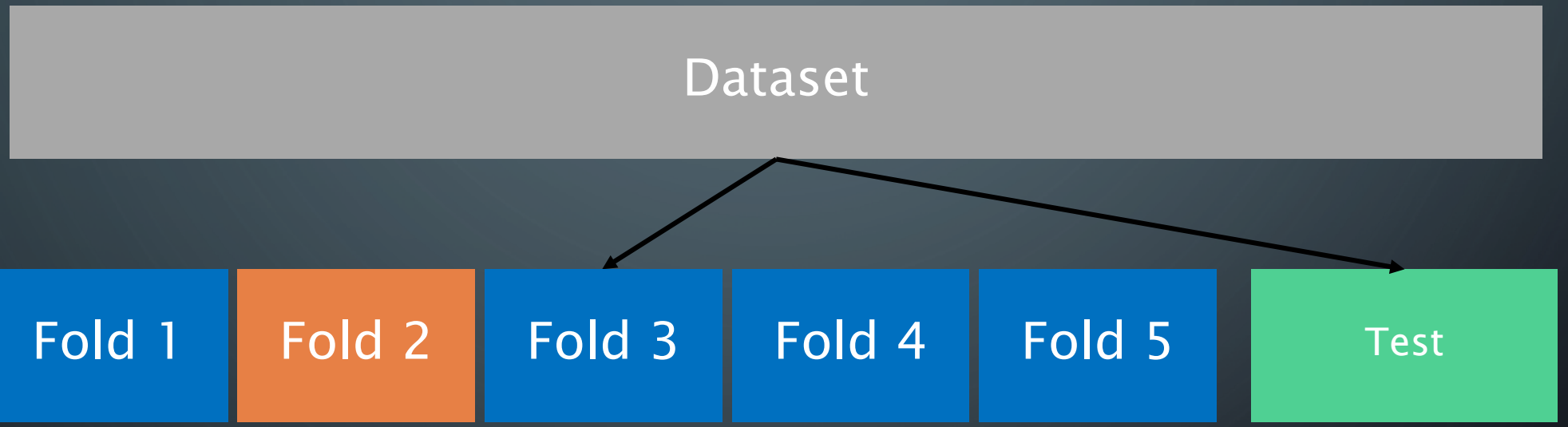
K-FOLD CROSS VALIDATION



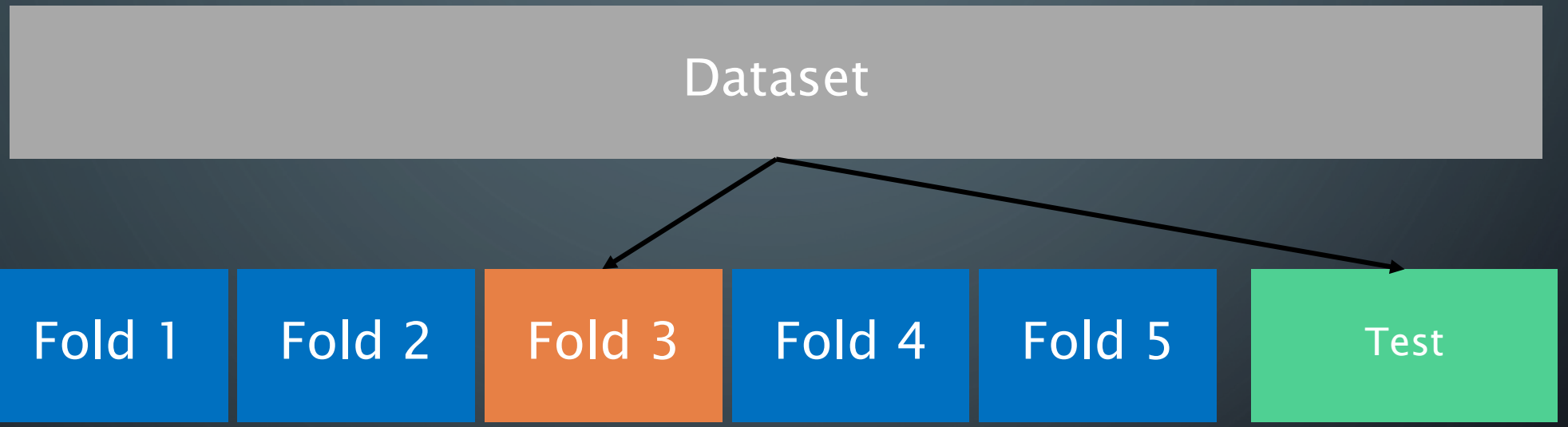
K-FOLD CROSS VALIDATION



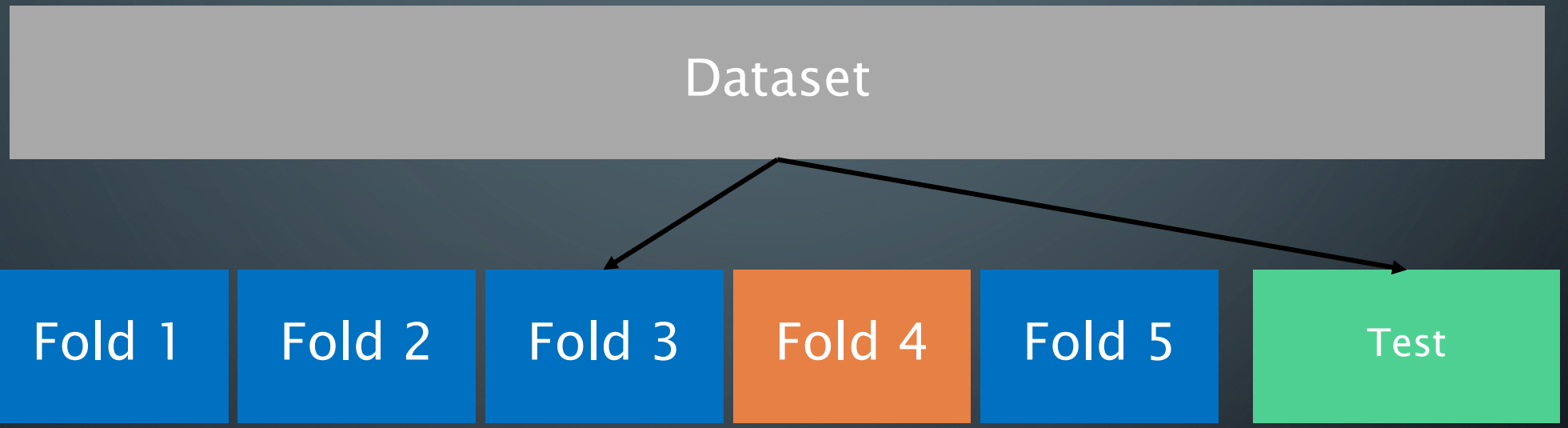
K-FOLD CROSS VALIDATION



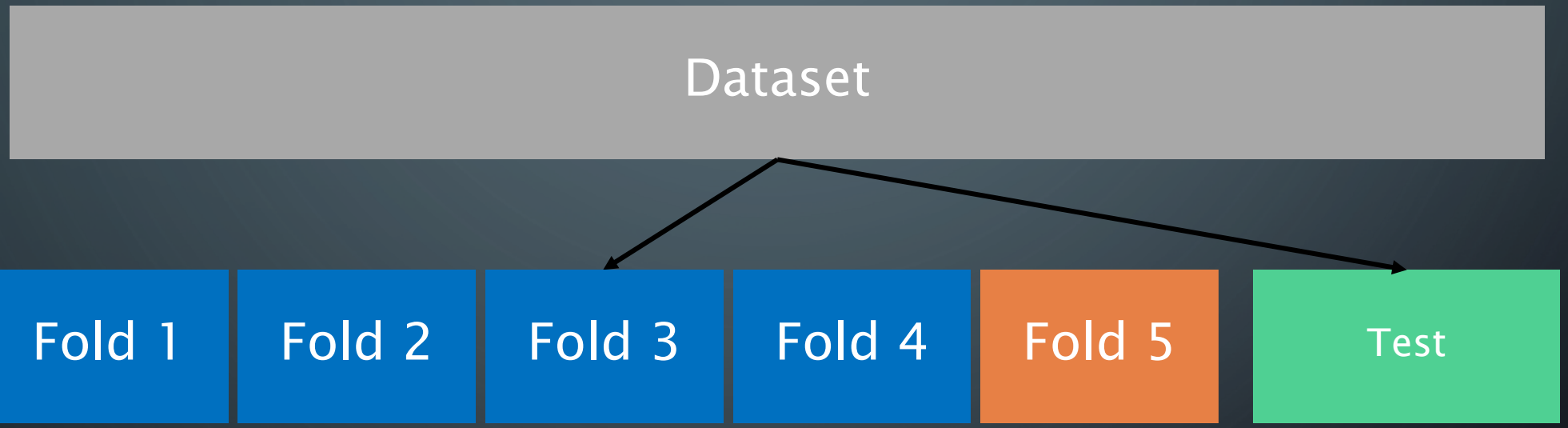
K-FOLD CROSS VALIDATION



K-FOLD CROSS VALIDATION



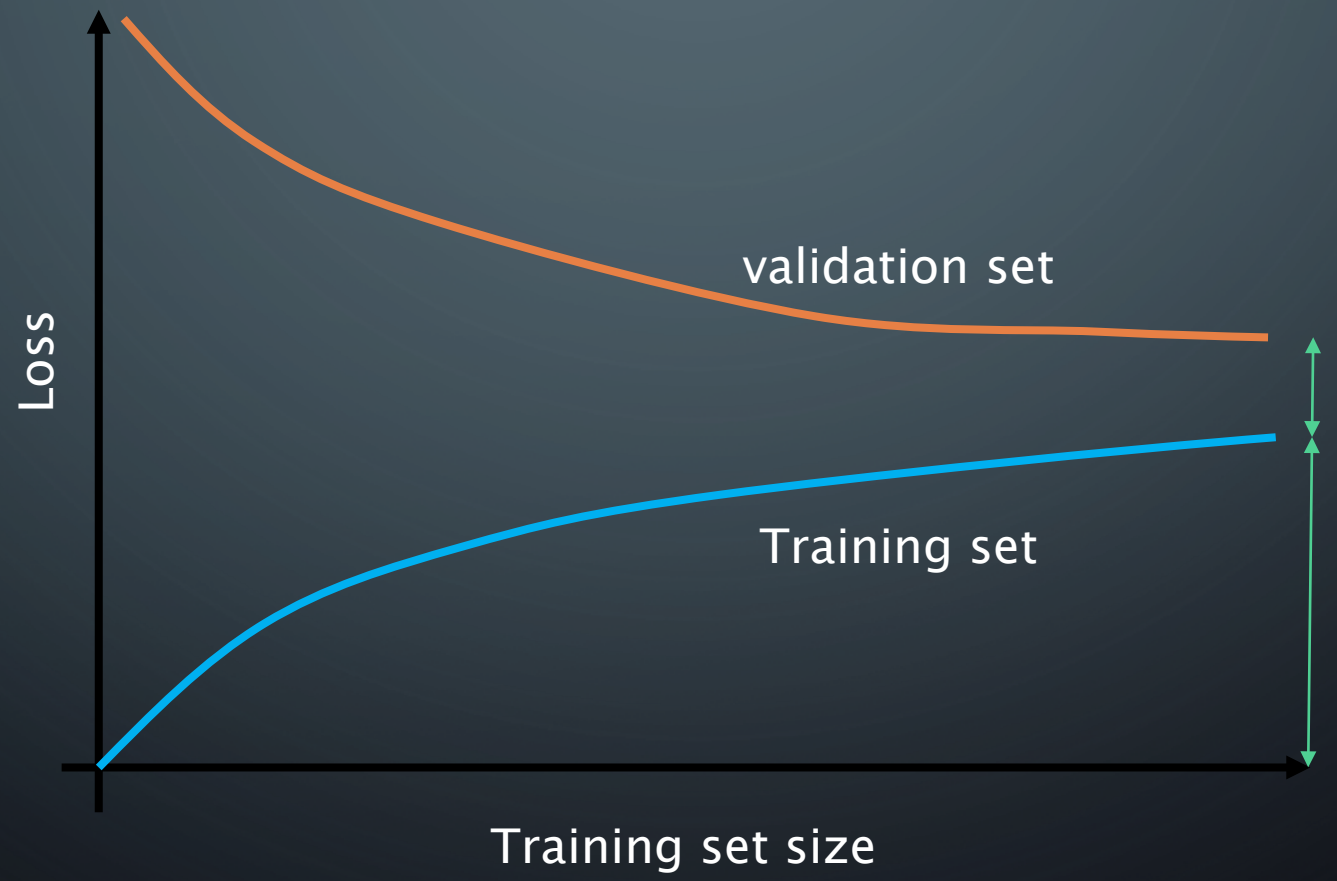
K-FOLD CROSS VALIDATION



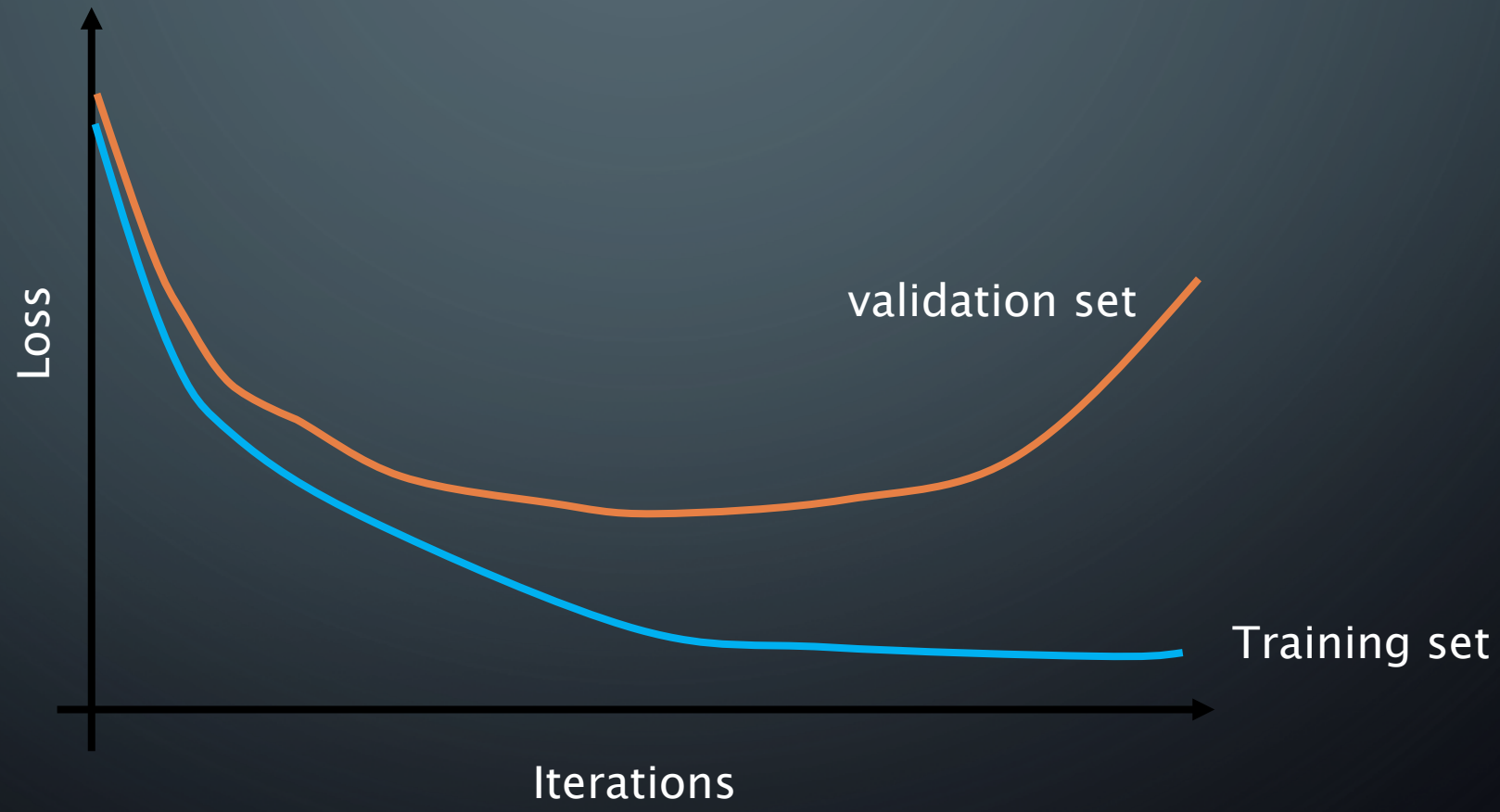
HYPER PARAMETER TUNING

- Grid search
- Random search
- Evolutionary optimization
- …

LEARNING CURVE (OLD)



LEARNING CURVE (NEW)



HANDLING UNDERFITTING

- Increase model capacity
  - Add polynomial features
- Increase features
  - Get more features
  - Use feature extraction
- Reduce regularization
- Use more powerful models
- …

HANDLING OVERFITTING

- Reduce model capacity
  - Reduce parameters
  - Reduce polynomial degree
  - Remove some features
- Apply regularization
- Increase data size
- Early stopping
- …

CLASSIFICATION PERFORMANCE METRIC

- Accuracy
- Confusion matrix
- Precision
- Recall
- F1-score
- ROC-AUC

PERFORMANCE METRIC (ACC)

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

$$Accuracy(ACC) = \frac{\#TP + \#TN}{\#Positive + \#Negative}$$

PERFORMANCE METRIC (PRECISION)

		Actual	
		Positive	Negative
Predicted	Positive	**True Positive**	**False Positive**
	Negative	**False Negative**	**True Negative**

$$Precision = \frac{\#TP}{\#TP + \#FP}$$

PERFORMANCE METRIC (RECALL)

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

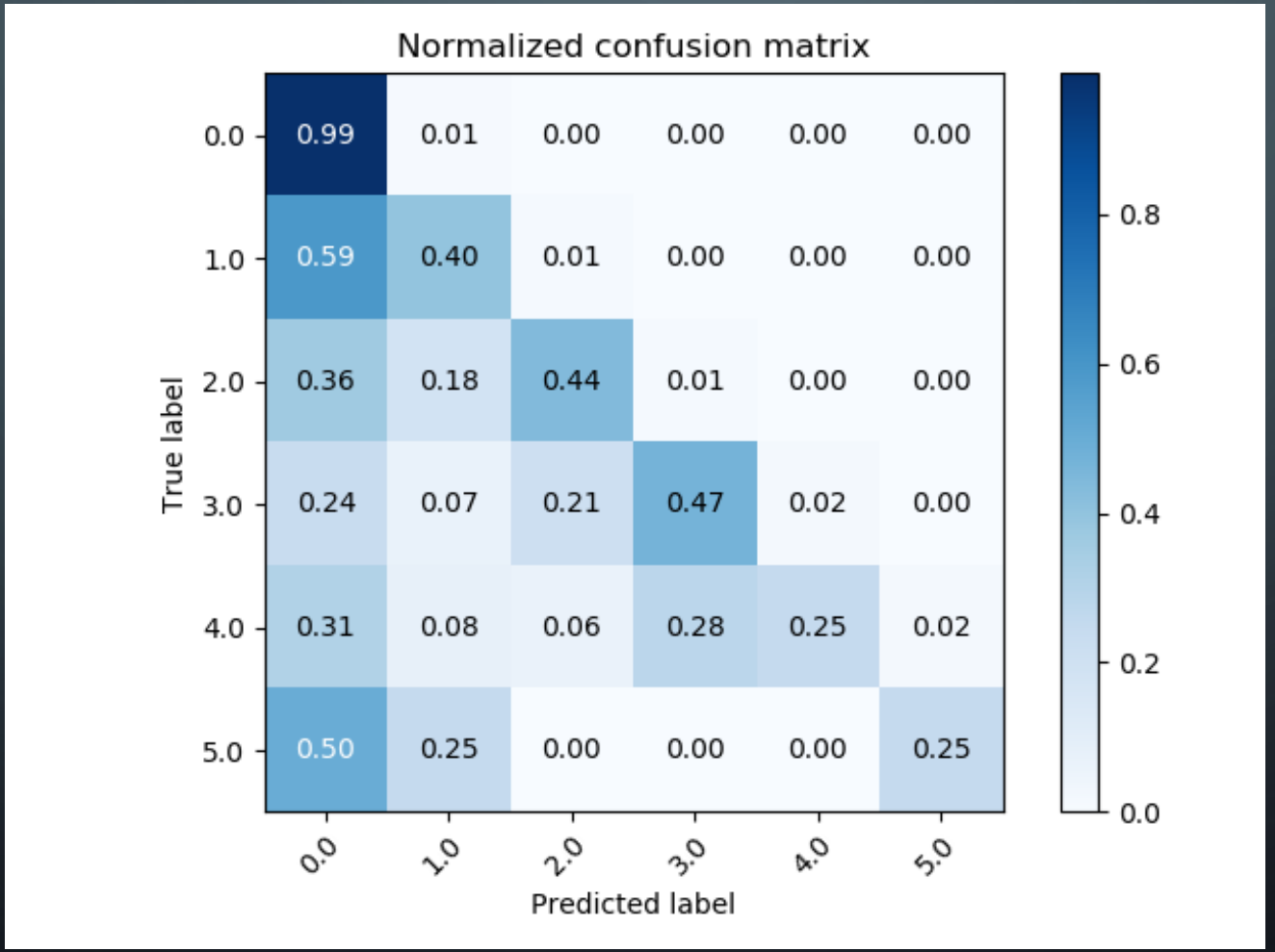
$$recall(TPR) = \frac{\#TP}{\#TP + \#FN}$$

PERFORMANCE METRIC (F1-SCORE)

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

$$F_1\ score = 2 * \frac{Precision\ * Recall}{Precision + Recall}$$

CONFUSION MATRIX



REGRESSION PERFORMANCE METRIC

- Mean Absolute Error (MAE)
- Mean Squared Error (MSE)
- Root Mean Absolute Error (RMAE)
- R-Squared
- Adjusted R-Squared

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|h_i - y_i|$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(h_i - y_i)^2$$

$$RMSE = \sqrt{MSE}$$

SEGMENTATION PERFORMANCE METRIC

- Pixel Accuracy
- Intersection-over-Union (IoU)
- Dice coefficient
- DICE and IoU are correlated

$$Dice = 2\frac{|X \cap Y|}{|X| + |Y|} \qquad IoU = \frac{|X \cap Y|}{|X \cup Y|}$$