CONFUSION MATRIX, ROC-AUC, AND BIAS-VARIANCE TRADEOFF

MACHINE LEARNING EVALUATION CONCEPTS

1. CONFUSION MATRIX

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

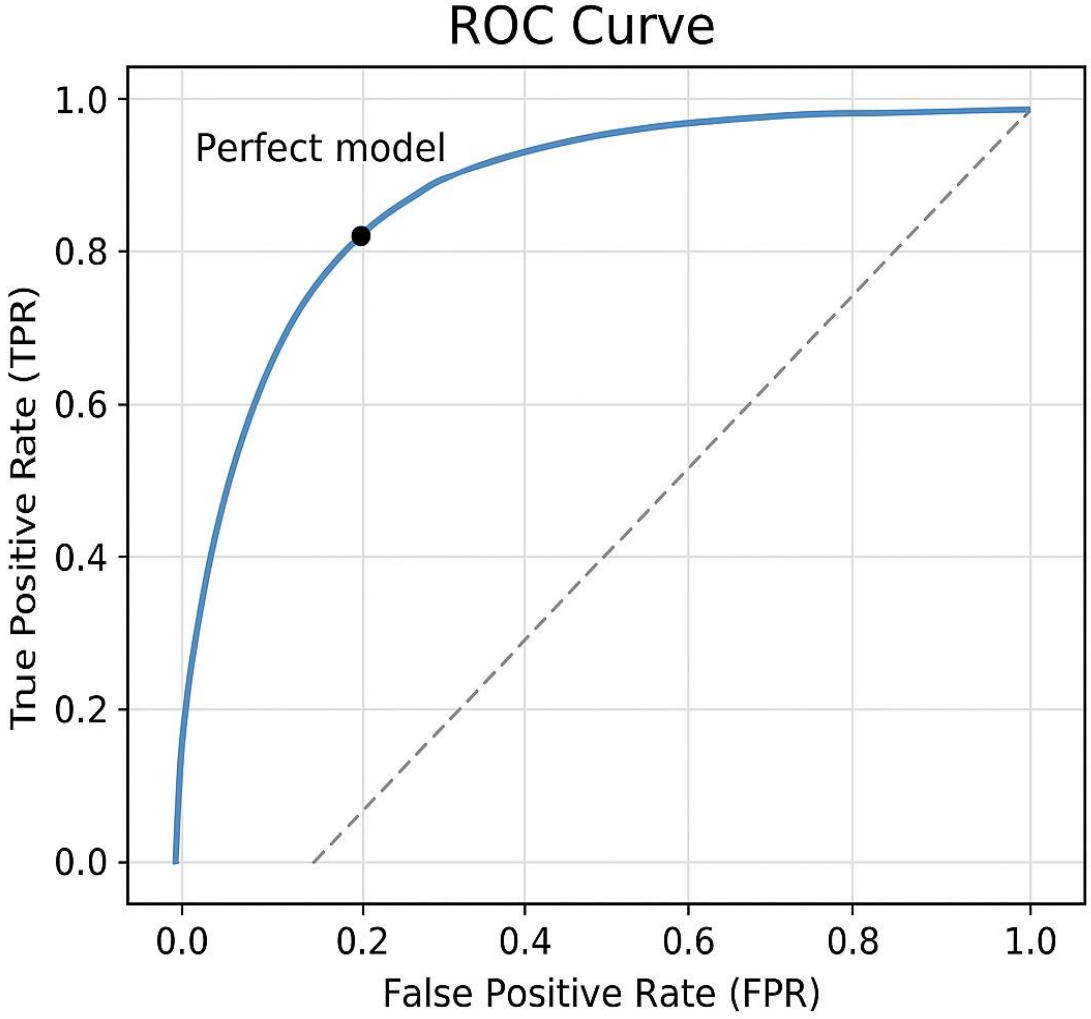
CONFUSION MATRIX METRICS

FROM THE CONFUSION MATRIX, WE CALCULATE:

- ACCURACY = (TP+TN)/(TP+TN+FP+FN)
- PRECISION = TP/(TP+FP)
- RECALL (SENSITIVITY) = TP/(TP+FN)
- SPECIFICITY = TN/(TN+FP)
- F1 SCORE=2*(PRECISION*RECALL)/(PRECISION+RECALL)

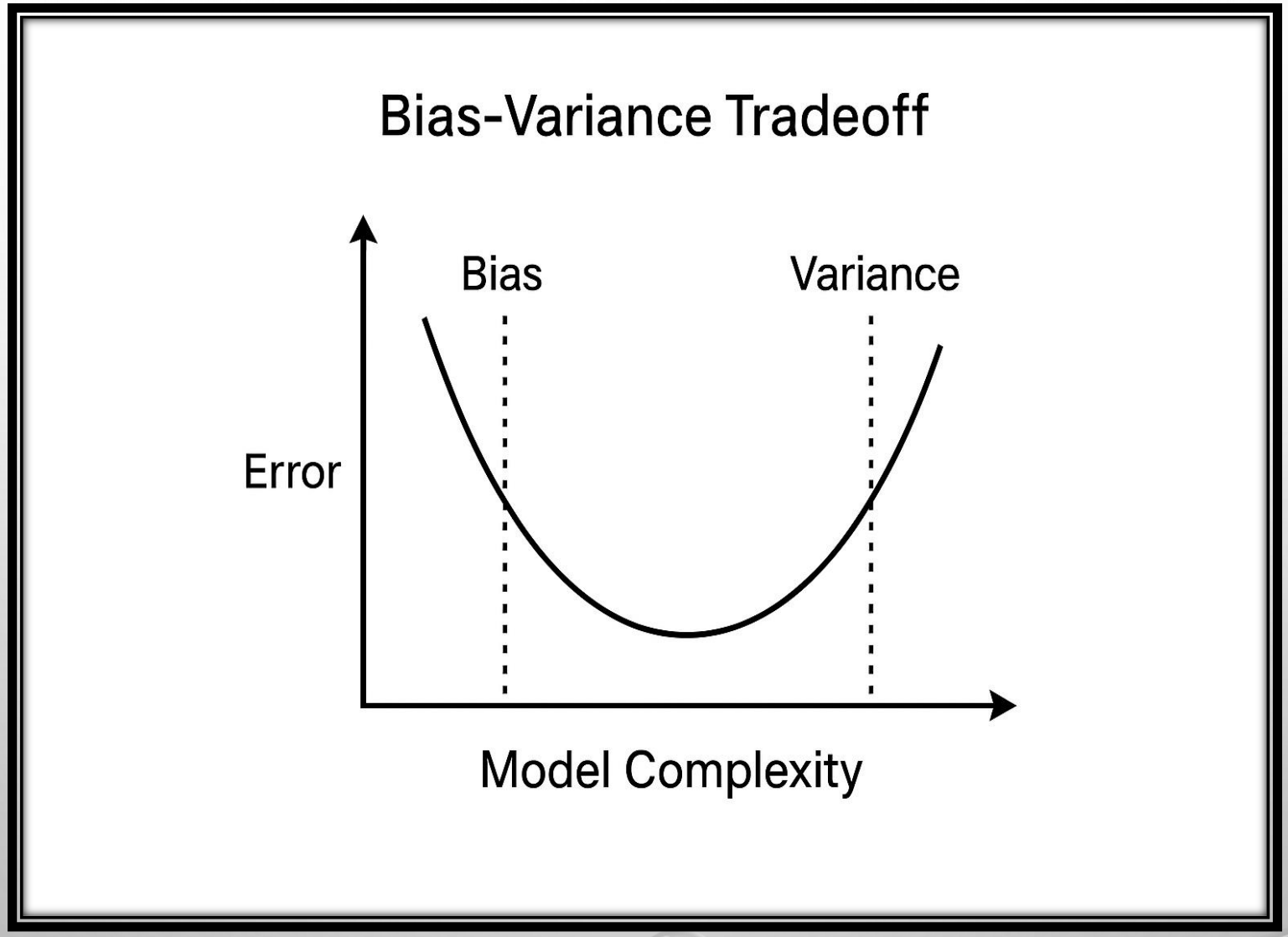
2. ROC CURVE AND AUC

- ROC CURVE PLOTS TRUE POSITIVE RATE (TPR) VS FALSE POSITIVE RATE (FPR).
- AUC MEASURES THE AREA UNDER THIS CURVE.
- AUC = 1 → PERFECT CLASSIFIER, AUC = 0.5 → RANDOM GUESSING.


ROC Curve
Perfect model
True Positive Rate (TPR)
False Positive Rate (FPR)
0.0
0.2
0.4
0.6
0.8
1.0

3. BIAS–VARIANCE TRADEOFF

- BIAS: ERROR DUE TO OVERSIMPLIFICATION (UNDERFITTING).
- VARIANCE: ERROR DUE TO TOO MUCH COMPLEXITY (OVERFITTING).
- THE BEST POINT IS THE TRADEOFF BETWEEN BIAS AND VARIANCE.


Bias-Variance Tradeoff
Bias
Variance
Error
Model Complexity