Metrics and Model Analysis

Confusion Matrix, MSE, MAE, Bias-variance tradeoff

COGNIBOT
AI meets Industry

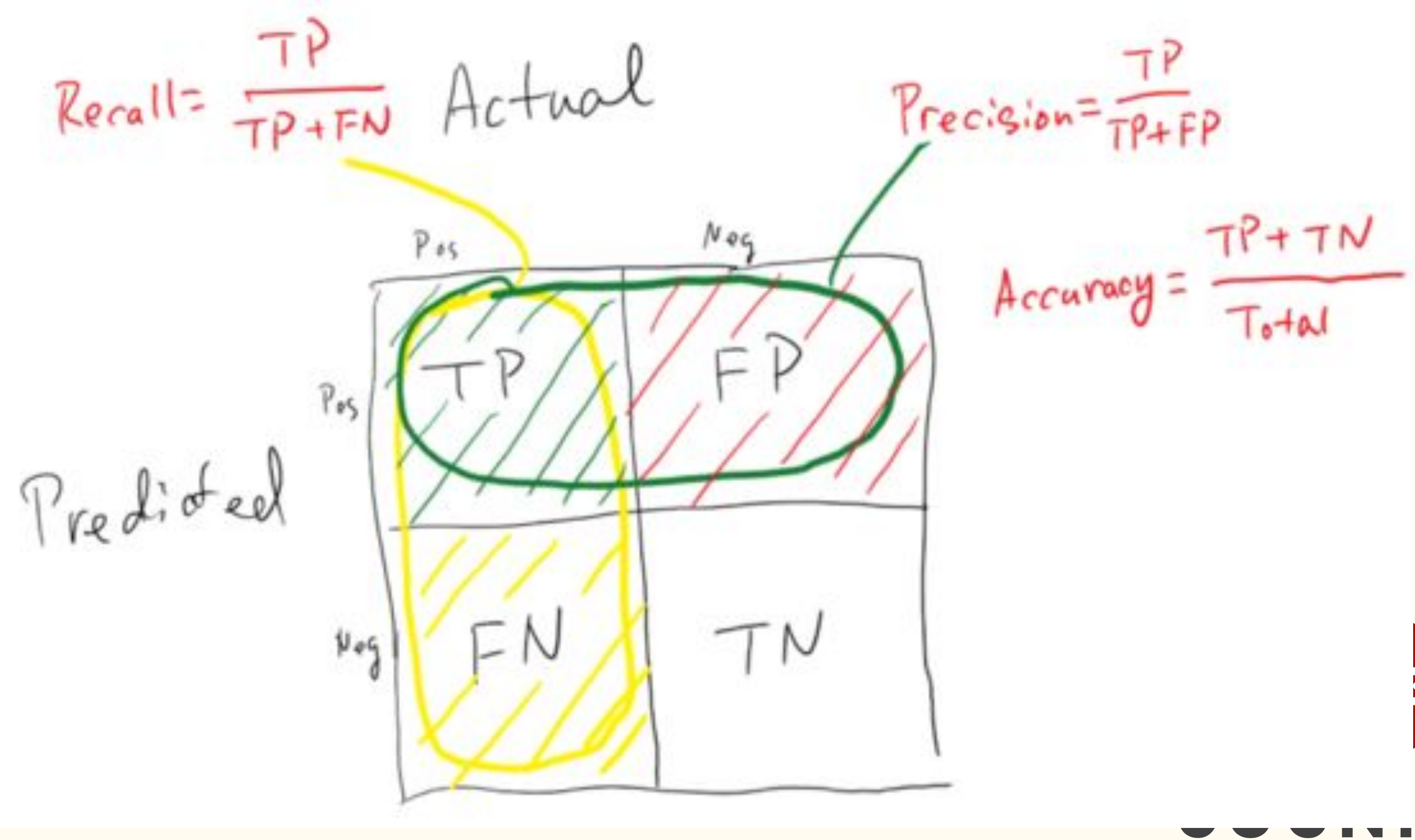
Classification problems

Confusion Matrix

COGNIBOT
AI meets Industry

- A confusion matrix is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known.

Predicted Value (predicted by the test)		Actual Value (as confirmed by experiment)	
		positives	negatives
	positives	**TP** True Positive	**FP** False Positive
	negatives	**FN** False Negative	**TN** True Negative



AI meets Industry

Problems with Accuracy

- Assumes equal cost for both kinds of error
- Accuracy alone doesn't tell the full story when you're working with a class-imbalanced data set

Example

True Positive (TP): • Reality: Malignant • ML model predicted: Malignant • **Number of TP results: 1**	**False Positive (FP):** • Reality: Benign • ML model predicted: Malignant • **Number of FP results: 1**
False Negative (FN): • Reality: Malignant • ML model predicted: Benign • **Number of FN results: 8**	**True Negative (TN):** • Reality: Benign • ML model predicted: Benign • **Number of TN results: 90**

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{1 + 90}{1 + 90 + 1 + 8} = 0.91$$

- Of the 91 benign tumors, the model correctly identifies 90 as benign. However, of the 9 malignant tumors, the model only correctly identifies 1 as malignant—a terrible outcome, as 8 out of 9 malignancies go undiagnosed!
- While 91% accuracy may seem good at first glance, another tumor-classifier model that always predicts benign would achieve the exact same accuracy (91/100 correct predictions) on our examples.

COGNIBOT
AI meets Industry

Better Alternatives:

- Precision: What proportion of positive identifications was actually correct?

True Positives (TPs): 1	False Positives (FPs): 1
False Negatives (FNs): 8	True Negatives (TNs): 90

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{1}{1+1} = 0.5$$

- Our model has a precision of 0.5—in other words, when it predicts a tumor is malignant, it is correct 50% of the time.

COGNIBOT
AI meets Industry

- Recall: What proportion of actual positives was identified correctly?

True Positives (TPs): 1	False Positives (FPs): 1
False Negatives (FNs): 8	True Negatives (TNs): 90

$$\text{Recall} = \frac{TP}{TP+FN} = \frac{1}{1+8} = 0.11$$

- Our model has a recall of 0.11—in other words, it correctly identifies 11% of all malignant tumors.

COGNIBOT
AI meets Industry

Regression problems

MAE, RMSE, MSE

COGNIBOT
AI meets Industry

Mean Absolute Error

- Less sensitive to outliers

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}.$$

COGNIBOT
AI meets Industry

Mean Squared Error

- Very sensitive to outliers

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2.$$

Root Mean Squared Error

- RMSD is the square root of the average of squared errors. The effect of each error on RMSD is proportional to the size of the squared error; thus larger errors have a disproportionately large effect on RMSD. Consequently, RMSD is sensitive to outliers.

$$RMSErrors = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y_i} - y_i)^2}{n}}$$

COGNIBOT
AI meets Industry

Bias-Variance tradeoff

COGNIBOT
AI meets Industry

Error due to bias

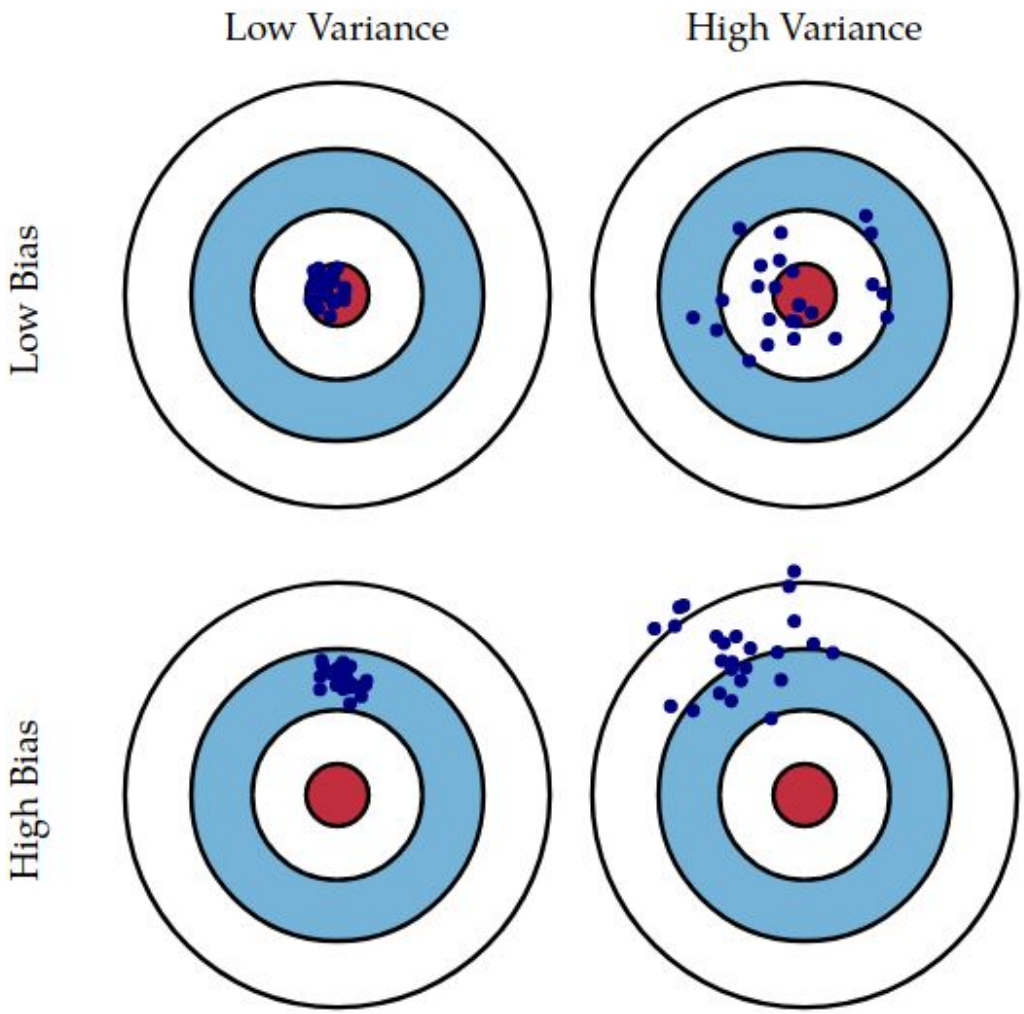
- The error due to bias is taken as the difference between the expected (or average) prediction of our model and the correct value which we are trying to predict.
- Bias measures how far off in general these models' predictions are from the correct value.

COGNIBOT
AI meets Industry

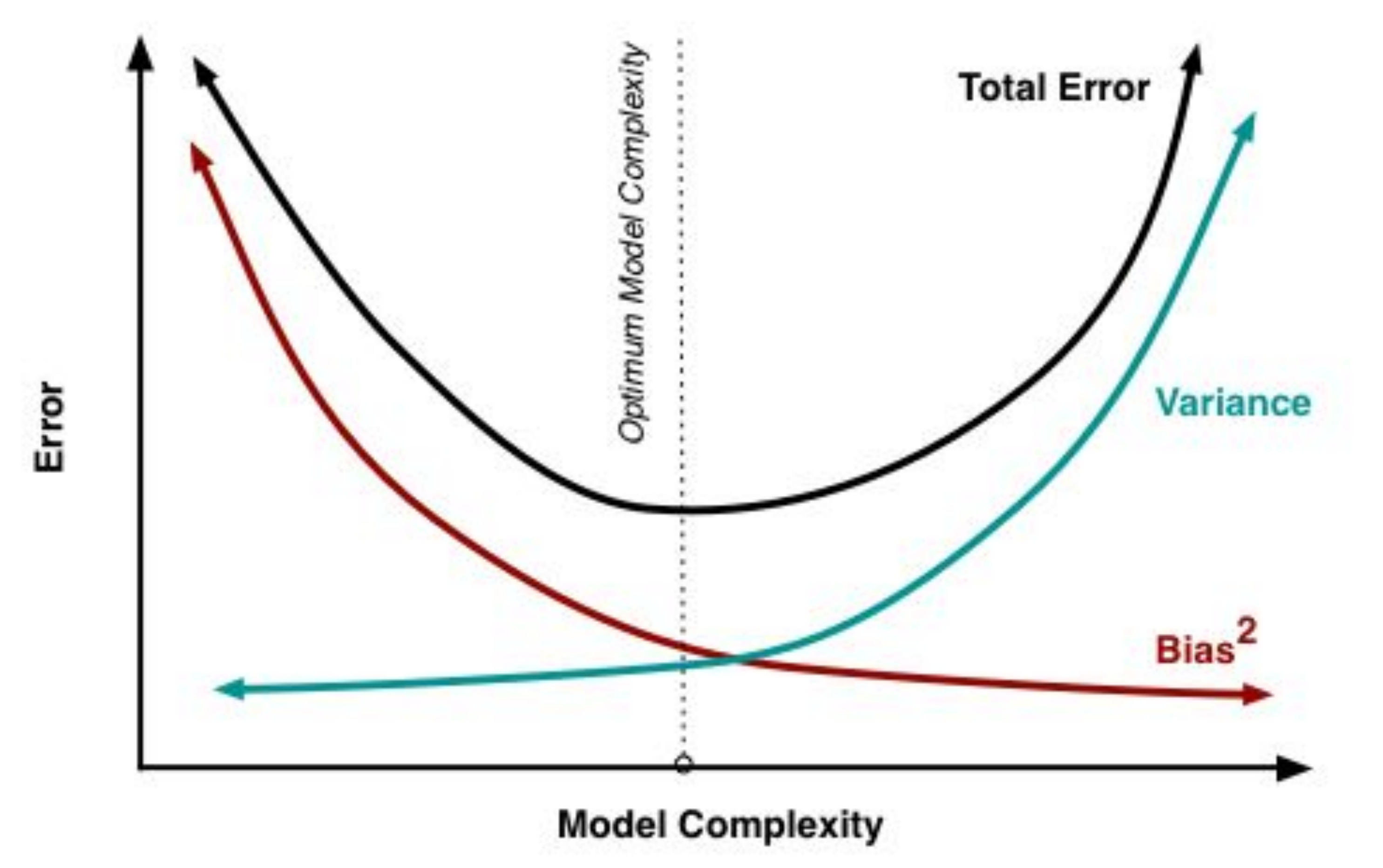
Error due to variance

- The error due to variance is taken as the variability of a model prediction for a given data point.
- The variance is how much the predictions for a given point vary between different realizations of the model.

COGNIBOT
AI meets Industry



AI meets Industry



IBOT
AI meets Industry