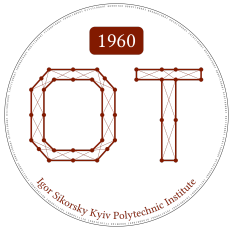


Computer Vision

Lecture 7: Evaluation Metrics for CV

Yuriy Kochura
iuriy.kochura@gmail.com
@y_kochura

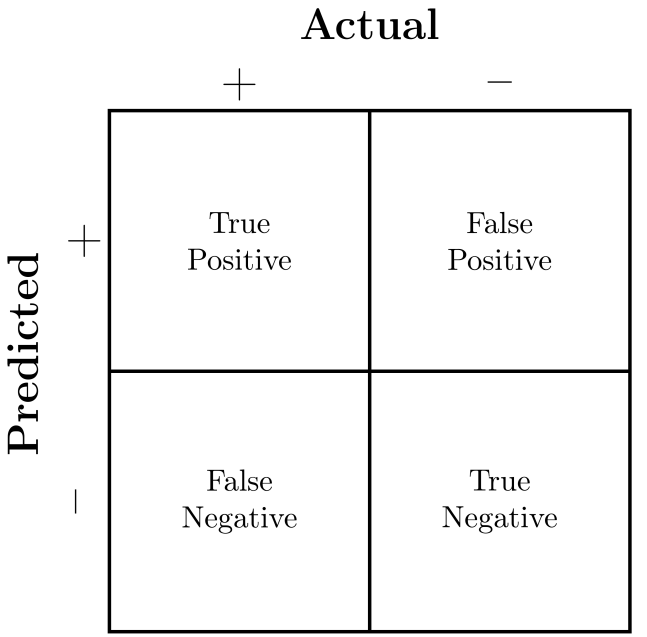
Today

How to measure a quality of a model, and easily compare different models on the same tasks?

- Classification
  - Accuracy
  - Precision
  - Recall
  - F1 Score
- Object Detection
  - Intersection over Union (IoU)
  - Average Precision (AP): the Area Under Curve (AUC)
  - Mean Average Precision (mAP)

Classification

Confusion Matrix

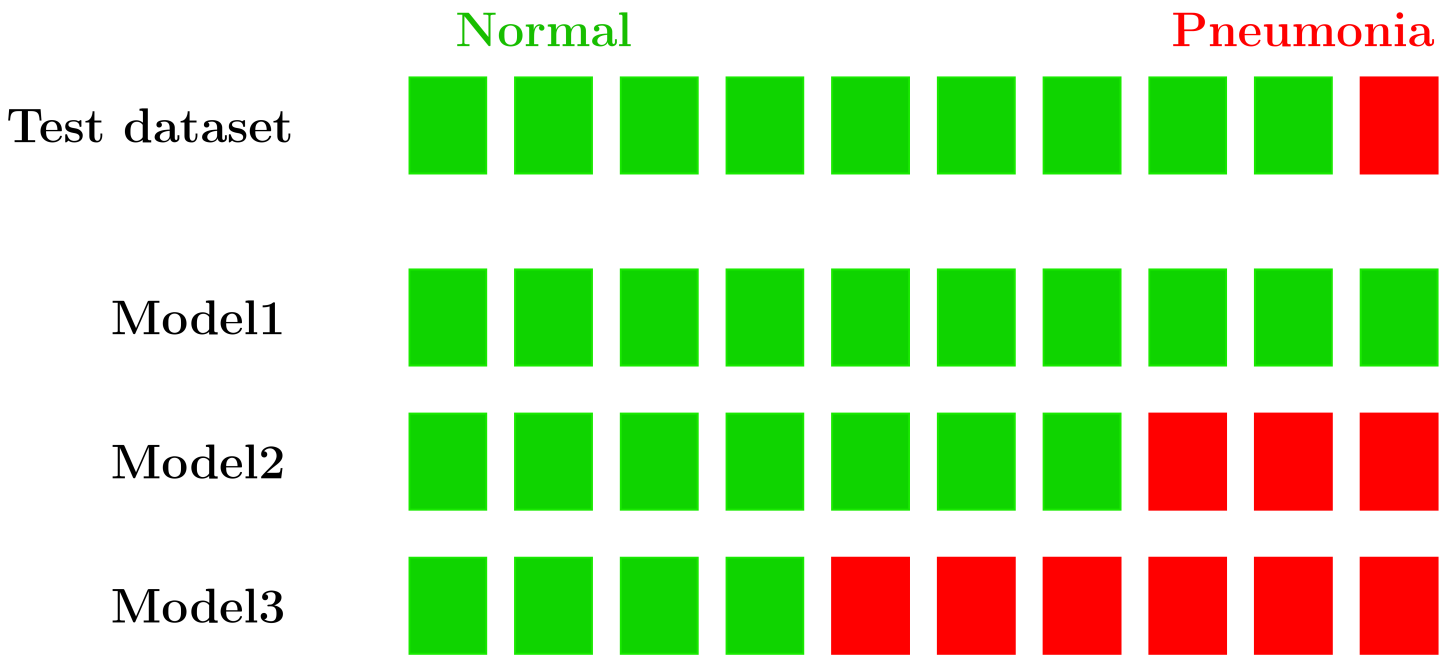


- True Positive (TP): Correctly identified as relevant
- True Negative (TN): Correctly identified as not relevant
- False Positive (FP): Incorrectly labeled as relevant
- False Negative (FN): Incorrectly labeled as not relevant

Example1: Identify Cats

Prediction	+	−	−	+	−	+
Image						
	True Positive	True Negative	False Negative	False Positive		

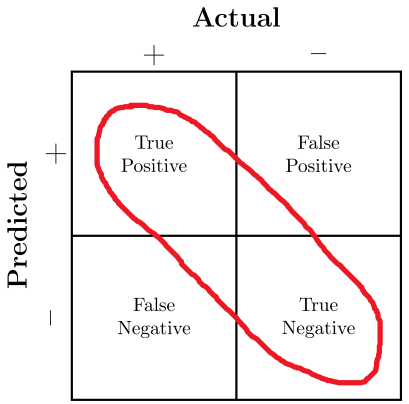
Example2: Identify pneumonia

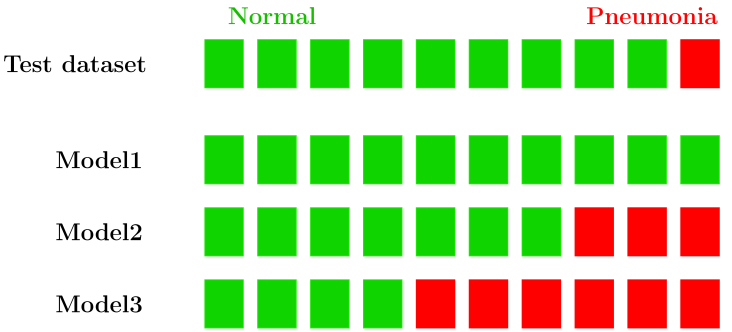


—

Credits: Andrew Ng and Kian Katanforoosh. CS230 Deep Learning.

Accuracy





$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

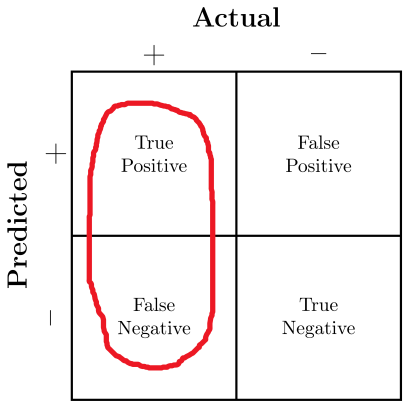
If we use accuracy as your evaluation metric, it seems that the best model is Model1:

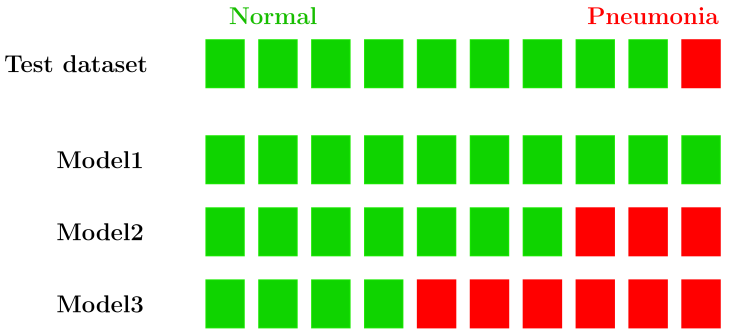
$$\text{Accuracy}(M_1) = \frac{9}{10} \qquad \text{Accuracy}(M_2) = \frac{8}{10} \qquad \text{Accuracy}(M_3) = \frac{5}{10}$$

—

Credits: Andrew Ng and Kian Katanforoosh. CS230 Deep Learning.

Recall





$$\text{Recall} = \frac{TP}{TP + FN}$$

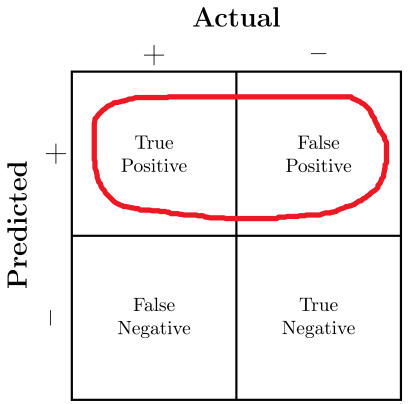
What is the recall for each model?

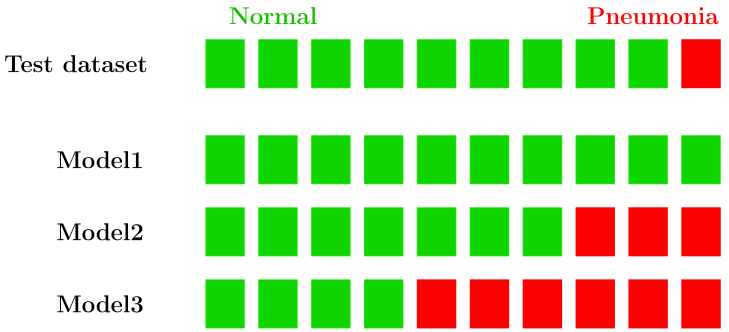
$$\text{Recall}(M_1) = \frac{0}{1} \qquad \text{Recall}(M_2) = \frac{1}{1} \qquad \text{Recall}(M_3) = \frac{1}{1}$$

—

Credits: Andrew Ng and Kian Katanforoosh. CS230 Deep Learning.

Precision





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

What is the precision for each model?

$$\text{Precision}(M_1) = \frac{0}{0} \qquad \text{Precision}(M_2) = \frac{1}{3} \qquad \text{Precision}(M_3) = \frac{1}{6}$$

—

Credits: Andrew Ng and Kian Katanforoosh. CS230 Deep Learning.

F1 score

$$F_1 = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

What is the F1 score for each model?

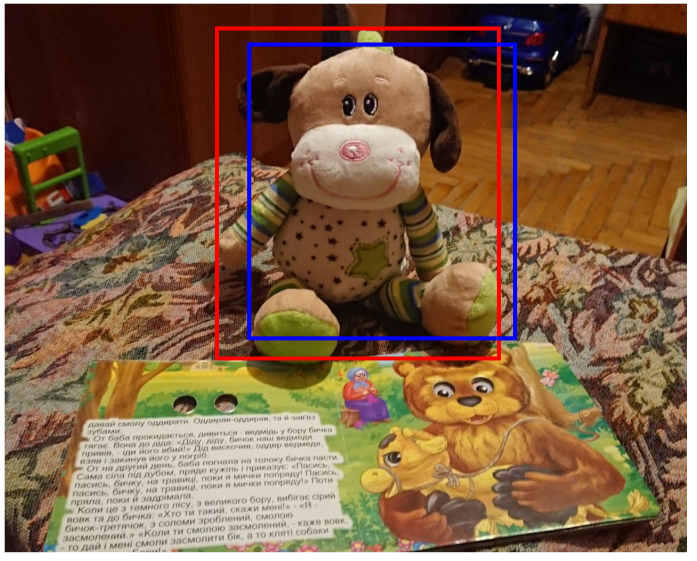
$$F_1(M_1) = 0 \qquad F_1(M_2) = \frac{1}{2} \qquad F_1(M_3) = \frac{2}{7}$$

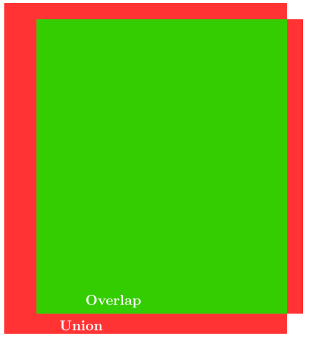
—

Credits: Andrew Ng and Kian Katanforoosh. CS230 Deep Learning.

Object Detection

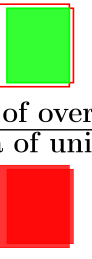
Intersection over Union (IoU)





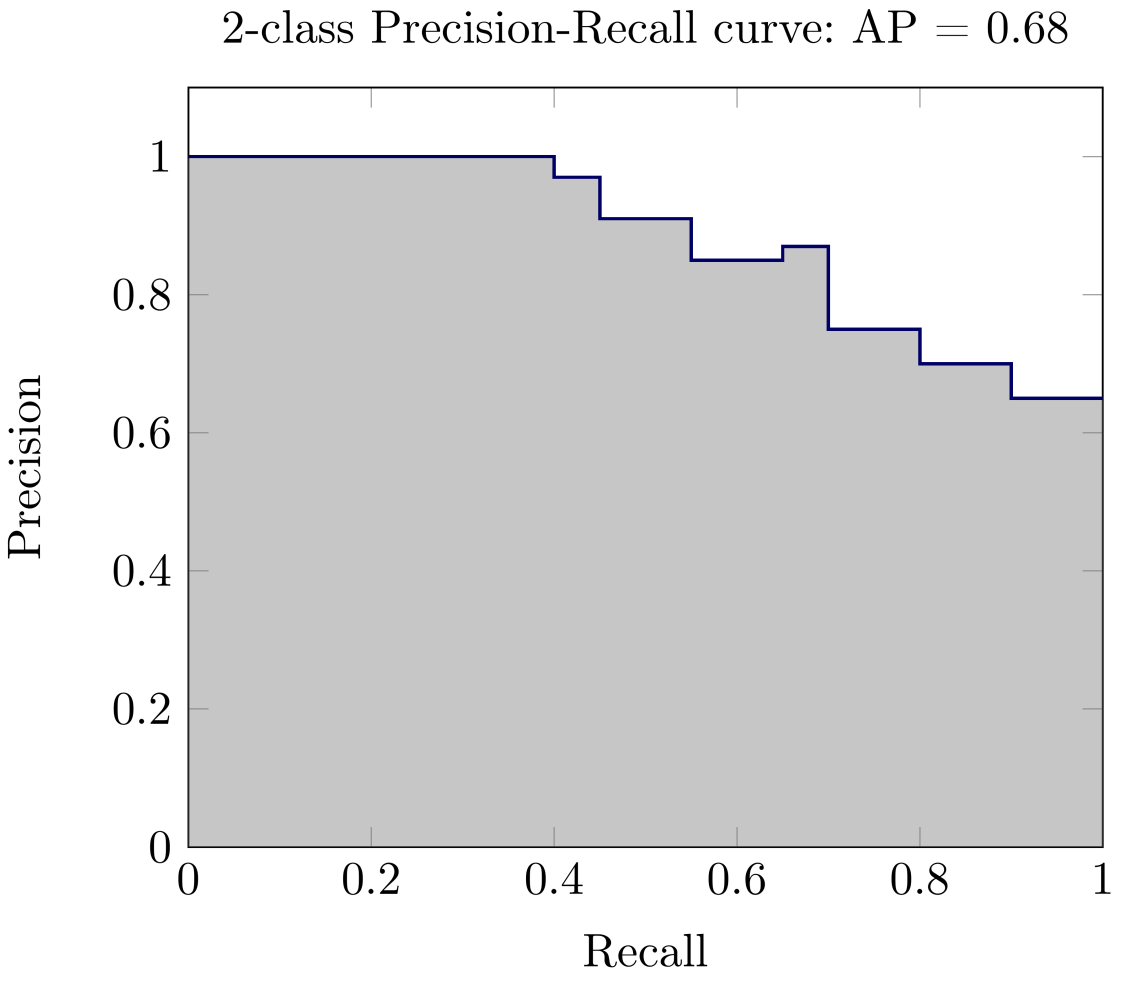
Ground Truth

Prediction



$$\text{IoU} = \frac{\text{area of overlap}}{\text{area of union}}$$

Average Precision (AP): the Area Under Curve (AUC)



Demo

Evaluation Metrics

The end