Chapter 3: Performance Measures in Classification

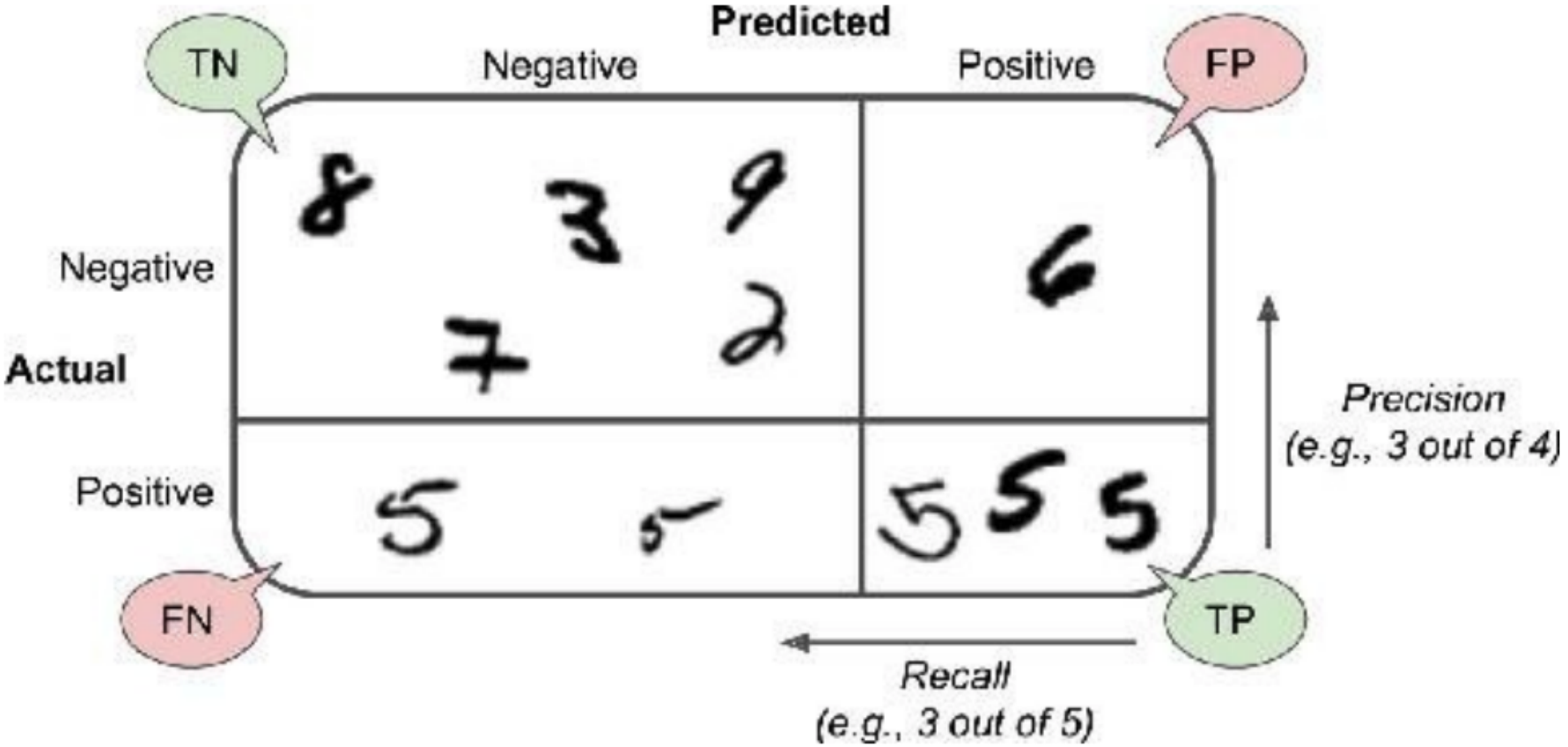
Dr. Xudong Liu
Assistant Professor
School of Computing
University of North Florida

Monday, 9/16/2019

Confusion Matrix

- Each row in a confusion matrix represents an *actual* class.
- Each colum in a confusion matrix represents a *predicted* class.
- Example (next slide): classify whether the digit in an image is 5.
- True negatives (TN): predicted negative examples that are actually negative.
- False positives (FP): predicted positive examples that are actually negative.
- False negatives (FN): predicted negative examples that are actually positive.
- True positives (TP): predicted positive examples that are actually positive.
- In Python, you may get it with *confusion_matrix* method.
- Can extend to multi-class classification problems.

Confusion Matrix



Precision

- Precision is a way to look at the accuracy of the positive predictions:

$$precision = \frac{|TP|}{|TP|+|FP|}$$

- For the previous example, precision is 75%.
- But precision can be 100% if the classifier only makes one positive prediction that is correct. This would NOT be useful.

Recall

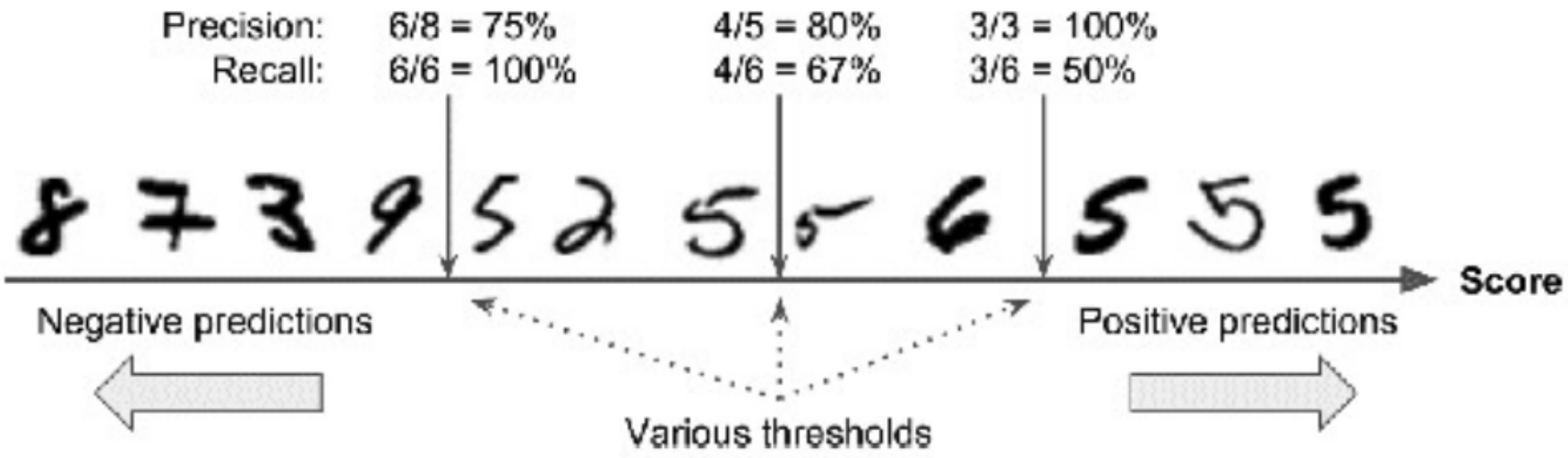
- Recall is a way to look at the percentage of positive examples predicted correctly:

$$recall = \frac{|TP|}{|TP|+|FN|}$$

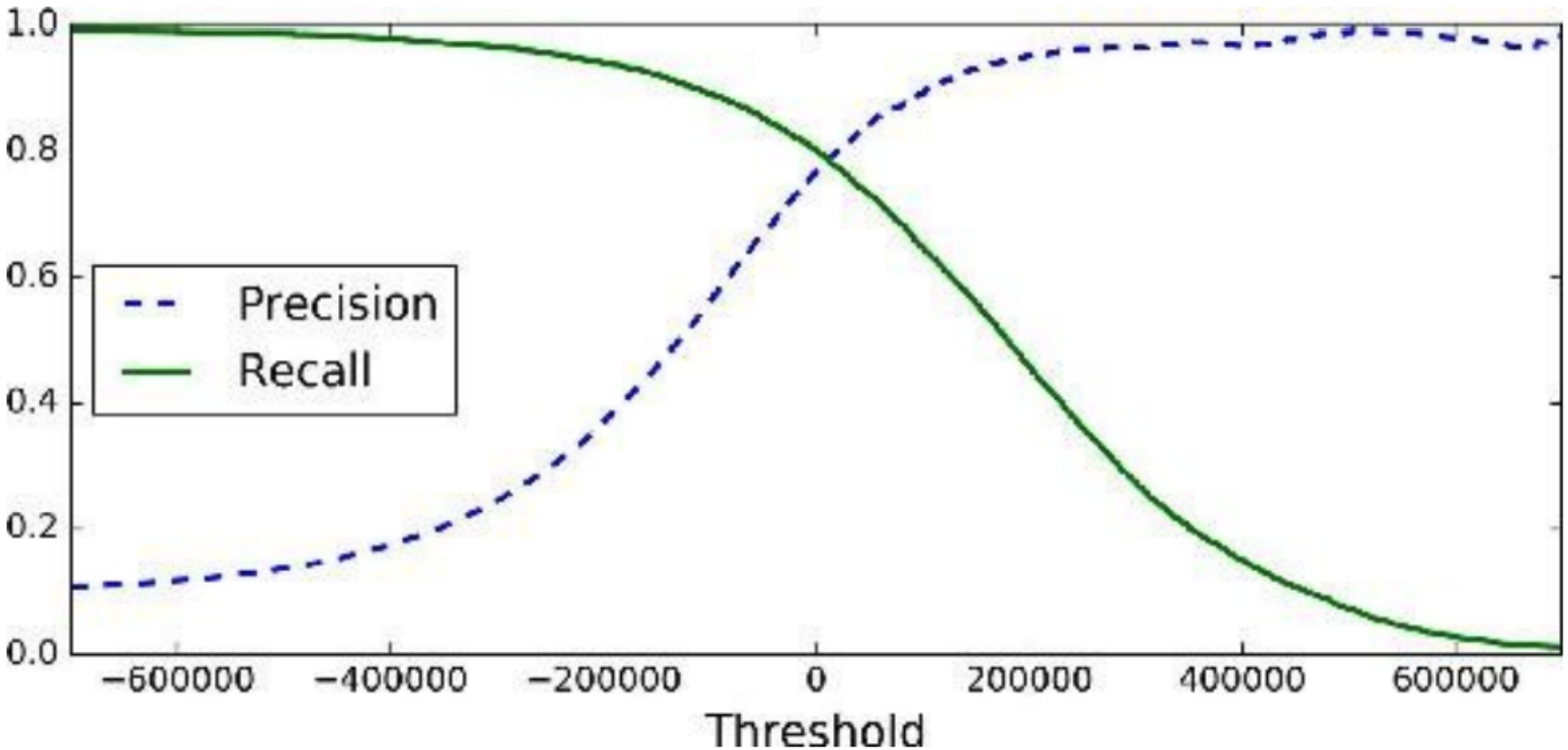
- For the previous example, precision is 60%.

Precision/Recall Trade-off

- Classification models, e.g., SGDClassifier, often predict based on a computed score of a given example.
- If the score is below a set threshold, the example is predict negative; otherwise, positive.
- For the previous setting, all examples are sorted based on their scores.



Precision/Recall Trade-off



F_1 Score

- The F_1 score is the harmonic mean of precision and recall:

$$F_1 = \frac{2}{\frac{1}{precision} + \frac{1}{recall}}$$

- Unlike regular mean, harmonic mean gives more weight to low values.
- Therefore, the classifier's F_1 score is only high if both recall and precision are high.