Evaluation Measures of an Algorithm

Zahoor Tanoli (PhD)
CUI, Attock

WHY IS ACCURACY NOT A GOOD METRIC TO EVALUATE CLASSIFIERS?

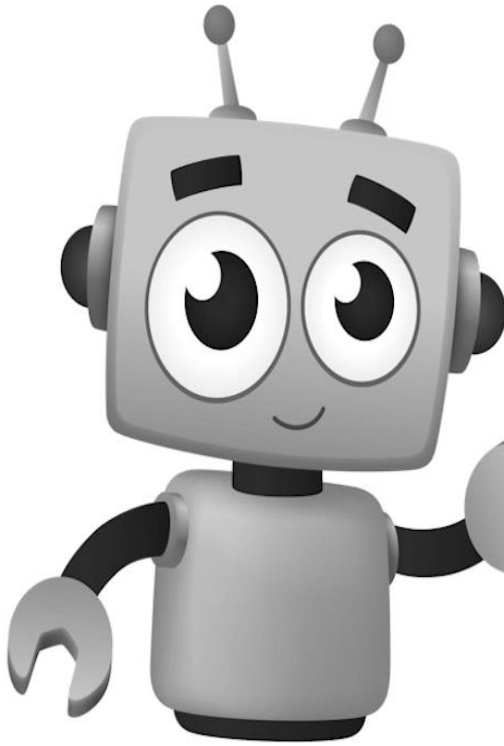
130 Cr =

D → 4 (130 cr - 4) MNIST → 6 1 [2] → True.
7 9

4/5 = 80%

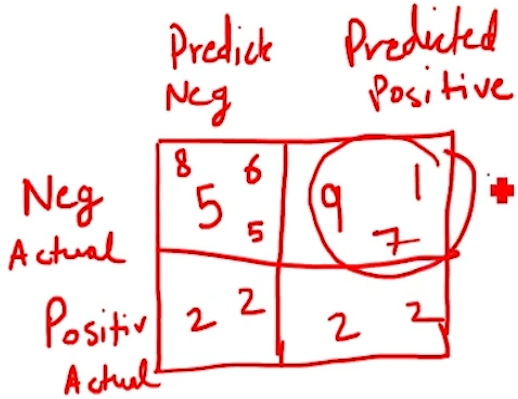
Dumb.

- We saw accuracy as a metric to evaluate classifiers
- Accuracy is not always the best way to evaluate classifiers (especially in skewed datasets)
- Evaluating a classifier is trickier than evaluating a regressor



WHAT IS A CONFUSION MATRIX?

- It is a much better way to evaluate classifiers
- How does it work?
- Just count the number of times instances of class A are classified (or confused) as class B

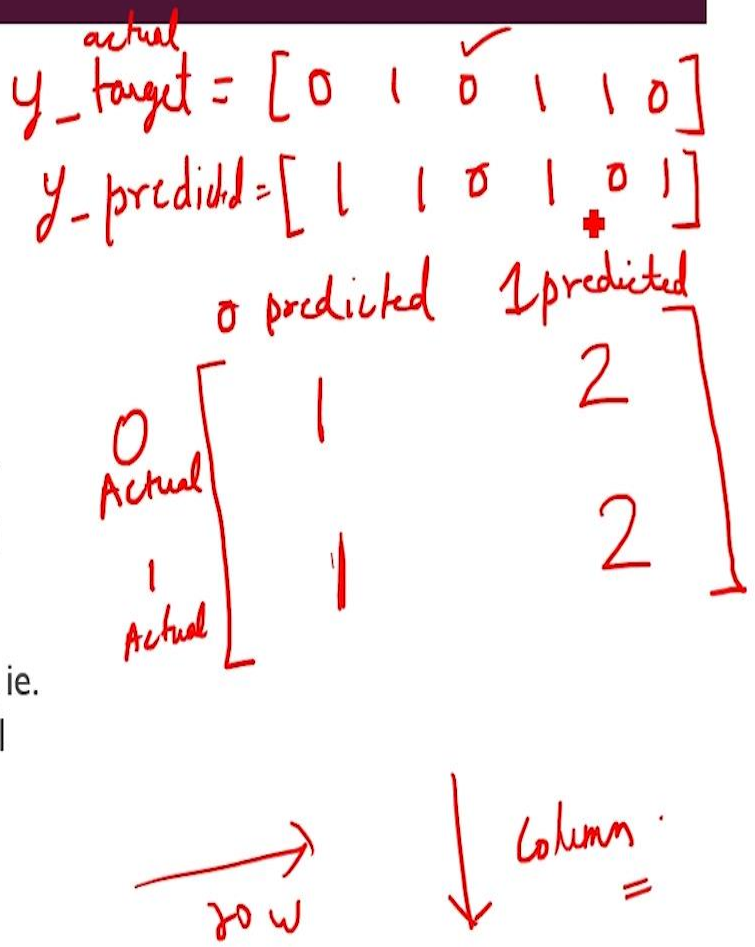


$$\begin{bmatrix} 4 & 3 \\ 2 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 71 & 0 \\ 0 & 22 \end{bmatrix}$$

HOW TO CREATE CONFUSION MATRIX?

- Get the set of predictions, so they can be compared to the actual targets.
- Each row in a confusion matrix represents an actual class, while each column repre- sents a predicted class
- First row is actual negative class & second row is actual positive class.
- First column is predicted negative class & second column is predicted positive class.
- Best classifier is the one having only true positives and true negatives, ie. confusion matrix would have nonzero values only on its main diagonal (top left to bottom right):



CONFUSION MATRIX

	Predicted negative classifier ke hisaab se 2 nahi hain ↓	Predicted positive classifier ke hisaab se 2 hain ↓
Actual negative 2 nahi hain →	7 9 7 5 1 TN	8 FP
Actual positive 2 hain →	2 2 FN	2 2 2 TP

Let us say we want to detect hand written 2s

CONFUSION MATRIX

	Predicted negative Classifier ke hisaab se 2 nahi hain ↓	Predicted positive Classifier ke hisaab se 2 hain ↓
Actual negative 2 nahi hain →	7 9 7 1 5 TN count = 5	8 FP Count = 1
Actual positive 2 hain →	2 2 FN Count = 2	2 Count = 3 2 2 TP

Let us say we want to detect hand written 2s

3/4

WHAT IS PRECISION?

10 → 2

- Precision in simple terms means – What percent of positive predictions made were correct?
- In Mathematical terms,

$$\text{Precision} = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

$$\frac{TP}{TP + FP}$$

WHAT IS RECALL?

- Recall in simple terms means – What percent of Actual positive values were correctly classified by your classifier?
- In Mathematical terms,

$$\text{Recall} = \frac{True\ Positives}{True\ Positives + False\ Negatives} = \frac{TP}{TP+FN}$$

PRECISION & RECALL FROM CONFUSION MATRIX

Rows = Actual Values ✓ *Columns = Predicted Values* ✓	**Predicted negative**	**Predicted positive**
Actual negative	7 9 7 1 5	8
Actual positive	2 2	2 2 2

Let us say we want to detect hand written 2s

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

$$\text{recall} = \frac{TP}{TP+fN}$$

F1- SCORE

99.9%

0.1%

- It is convenient to combine the performance of a classifier (precision and recall) into a single metric called the F1 score.
- F1-Score is the harmonic mean of precision and recall
- In Mathematical terms,

$$\text{F1-Score} = \frac{2\ X\ Precison\ X\ Recall}{Precision + Recall}$$

10,000

$$f_1 = \frac{1}{\frac{\frac{1}{prec} + \frac{1}{recall}}{2}}$$

Quality

Ambiance

Precision

f_1 — Recall

Thank You