Recall vs Precision

DATE: PAGE:

To make myself able to remember their meaning without thinking about true positive / false positive / false negative jargon, I conceptualize them as follows:

Imagine that, your girlfriend gave you a birthday surprise every year in last 10 years.

However, one day, your girlfriend asks you:

"Sweetie, do you remember all birthday surprises from me?"

This simple question makes your life in danger.

To extend your life, you need to recall all 10 surprising events from your memory.

So, recall is the ratio of a number of events you can correctly recall to a " " all correct events.

If you can recall all 10 events correctly, then, your recall ratio is 1.0 (100%). If you can recall 7 events correctly, your recall ratio is 0.7 (70%).

Now, it's easier to map the word recall to real life usage of that word.

However, you might be wrong in some answers.

For example, your answer 15 times, 10 events are correct and 5 events are wrong. This means

DATE: PAGE:

you can recall all events but it's not so precise

So, precision is the ratio of events you can correctly recall (events that your machine correctly predicts) to a number all events you recall (mix of correct and wrong recalls). In other words, it is how precise your recall is.

From the previous example (10 real events, 15 answers : 10 correct answers, 5 wrong answers), you get 100% recall but your precision is only 66.67% (10/15).

Therefore, if a machine learning algorithm is good at recall, it doesn't mean that algorithm is good at precision. That's why we also need F1 score which is the (harmonic) mean of recall & precision to evaluate an algorithm.

$$F = \frac{2PR}{(P+R)}$$

NOTE :

A number of events that you can correctly recall = True positive (they're correct and you recall them)

(predicted them as positive

A number of all correct events = True positive (they're correct and you recall them) + False negative (they're correct but you don't recall them)

you predicted them as negative but in fact they were positive

A number of all events you recall = True positive (they're correct and you recall them) + False " (they're not correct but you recall them)

- recall = TP/(TP+FN) Precision = TP/(TP+FP)