Explaining Accuracy, Precision and Recall with a Glass of Water

Sergey Medvedev
Jun 10 · 3 min read

Accuracy, **precision** and **recall** are three basic scores to measure the performance of a model in machine learning classification cases. Accuracy is a concept that is rather easy to grasp: take the sum of all *correct* predictions and divide them by the total number of predictions, *correct* and *wrong* ones. Naturally, if there are no wrong predictions, correct predictions devided by correct predictions will give us 1. Perfect accuracy!

A problem arises if the data is imbalanced which is often the case with classification problems. Say you have a database on credit card transactions with 95% of transactions being good ones and 5% being fraudulent. You've designed a model which correctly

predicts all 95% of the lawful transactions, but none of the bad ones. Following the formula, the model would have 95% accuracy level which is a very decent result. Unfortunately, that wouldn't prove your model to be good enough since it failed to detect any fraud. In those classification cases where the data is significantly imbalanced, accuracy score can be very misleading.

That is where precision and recall scores step in. Precision and recall formulas are as follows:

$$Precision = \frac{True\ Positive}{True\ Positive\ +\ False\ Positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive\ +\ False\ Negative}$$

While the variables in the formulas should not cause any problems, the ideas behind precision and recall can be harder to grasp at the beginning. It also does not help that precision and recall are measured separately for each class. But let me explain them to you with the following metaphor.
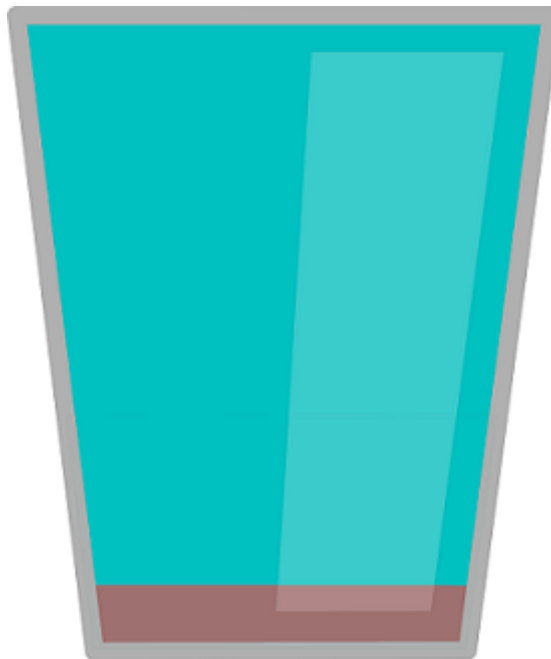
Imagine you have a glass of water. **Precision** will show you how much dirt there is in the water, i.e. how much there is of the substance that should not be there. It is as

simple as that — precision is about purity of the water in the glass and how much of an alien substance there is. The clearer the water — the higher the precision score will be. If your model claims something to be water that is not, that will lower the precision score.

Recall is a completely different story. It is not about purity, it is about amount. Imagine you've asked for a glass of water. They've brought it to you but it is not full. The question is how much water should be in the glass but is not there, i.e. how much water has been spilt — that is what the recall measures. The higher the recall is, the less of water has been lost on a way to you. If the recall equals 1, it means the glass is full — your model has predicted correctly *all* the records of a particular class, nothing has been overlooked, nothing's been left out of the glass.

To sum it up:

Precision measures how much dirt your model brings into the glass of water. **Recall** measures how much of water your model spills. It is as simple as that!

Machine Learning   Precision   Recall   Accuracy   F1

About   Help   Legal