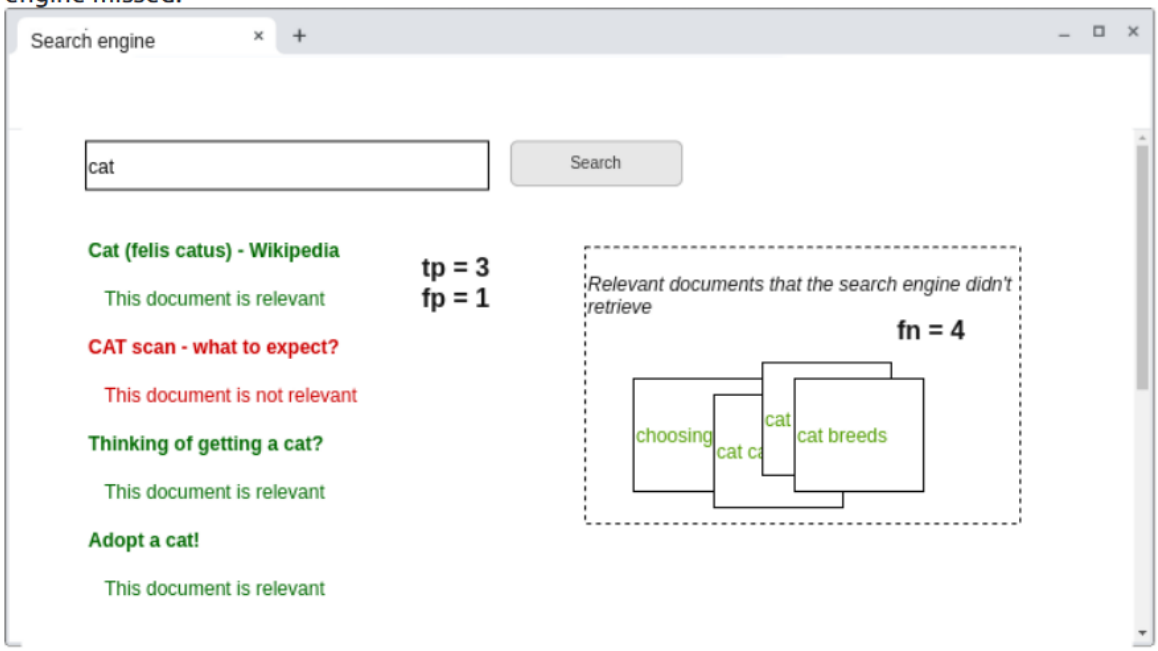
Understanding Precision and Recall with real world Examples

Example Calculation of Precision and Recall #1: Search engine

Imagine that you are searching for information about cats on your favorite search engine. You type *'cat'* into the search bar.

The search engine finds four web pages for you. Three pages are about cats, the topic of interest, and one page is about something entirely different, and the search engine gave it to you by mistake. In addition, there are four relevant documents on the internet, which the search engine missed.



In this case we have three true positives, so *tp* = 3. There is one false positive, *fp* = 1. And there are four false negatives, so *fn* = 4. Note that to calculate precision and recall, we do not need to know the total number of true negatives (the irrelevant documents which were not retrieved).

The precision is given by

$$\begin{aligned} \text{precision} &= \frac{\text{tp}}{\text{tp} + \text{fp}} \\ &= \frac{3}{3+1} \\ &= 0.75 \end{aligned}$$

and the recall is

$$\begin{aligned} \text{recall} &= \frac{\text{tp}}{\text{tp} + \text{fn}} \\ &= \frac{3}{3+4} \\ &= 0.43 \end{aligned}$$

Example Calculation of Precision and Recall #2: Disease diagnosis

Suppose we have a medical test which is able to identify patients with a certain disease. We test 20 patients and the test identifies 8 of them as having the disease. Of the 8 identified by the test, 5 actually had the disease (true positives), while the other 3 did not (false positives). We later find out that the test missed 4 additional patients who turned out to really have the disease (false negatives).

We can represent the 20 patients in the following confusion matrix:

		True state of patient's health	
		Disease	No disease
Test result	Alert	5	3
	No alert	4	8

The relevant values for calculating precision and recall are *tp* = 5, *fp* = 3, and *fn* = 4. Putting these values into the formulae for precision and recall, we obtain:

$$\begin{aligned}\text{precision} &= \frac{\text{tp}}{\text{tp} + \text{fp}} \\ &= \frac{5}{5 + 3} \\ &= 0.625\end{aligned}$$

$$\begin{aligned}\text{recall} &= \frac{\text{tp}}{\text{tp} + \text{fn}} \\ &= \frac{5}{5 + 4} \\ &= 0.56\end{aligned}$$

Applications of Precision and Recall in Information Retrieval

Precision and recall are best known for their use in evaluating search engines and other information retrieval systems.

Search engines must index large numbers of documents, and display a small number of relevant results to a user on demand. It is important for the user experience to ensure that both all relevant results are identified, and that as few as possible irrelevant documents are displayed to the user. For this reason, precision and recall are the natural choice for quantifying the performance of a search engine, with some small modifications.

Over 90% of users do not look past the first page of results. This means that the results on the second and third pages are not very relevant for evaluating a search engine in practice. For this reason, rather than calculating the standard precision and recall, we often calculate the precision for the first 10 results and call this precision @ 10. This allows us to have a measure of the precision that is more relevant to the user experience, for a user who is unlikely to look past the first page. Generalizing this, the precision for the first *k* results is called the precision @ *k*.

In fact, search engine overall performance is often expressed as mean average precision, which is the average of precision @ *k*, for a number of *k* values, and for a large set of search queries. This allows an evaluation of the search precision considering a variety of different user queries, and the possibility of users remaining on the first results page, vs scrolling through to the subsequent results pages.

Reference: https://deepai.org/machine-learning-glossary-and-terms/precision-and-recall