IN5400 Week 09: average precision

Alex

March 4, 2021

[The following notes are compiled from various sources such as textbooks, lecture materials, Web resources and are shared for academic purposes only, intended for use by students registered for a specific course. In the interest of brevity, every source is not cited. The compiler of these notes gratefully acknowledges all such sources.]

1 Understanding average precision

Consider the average precision score. It is a measure for ranking quality (higher is better).

- It assumes you have two classes $y_i \in \{0, 1\}$. You are predicting on a set of N samples $(x_i, y_i)_{i=1}^N$. Suppose $+1$ (from $y_i \in \{0, 1\}$) denotes the relevant class. Suppose $s(x_i)$ is a prediction score for class $+1$.
- Furthermore, in order to compute mean average precision, we assume that the samples (x_i, y_i) **are already sorted according to the prediction score $s(x_i)$ in descending order**, such that the highest scoring sample comes first.

  In information retrieval context it would mean that we rank the documents x_i according to a relevance prediction, so that the document which is predicted as most relevant, comes first.
- Next, we define to be the precision at k (P@k) as the precision for the top-k documents according:

$$P@k = \frac{1}{k} \sum_{i=1, \text{ sorted indices!!}}^{k} y_i$$

- This allows to define average precision as:

$$\begin{aligned}
AP &= \frac{1}{R}\sum_{k=1}^{N} 1[y_k == +1] P@k \\
&= \frac{1}{R}\sum_{k=1}^{N} 1[y_k == +1] \sum_{i=1,\text{ sorted !!}}^{k} \frac{1}{k} y_i \\
R &= \sum_{k=1}^{N} 1[y_k == +1]
\end{aligned}$$

Note that R is the number of samples on the dataset $(x_i, y_i)_{i=1}^N$ which are relevant ($y_i = +1$).

1.1 Questions:

- Suppose you have 11 total number of samples, 3 of them have $y_i = +1$ and those three come first in the ranking. what is your average precision?
- Suppose you have 11 total number of samples, 3 of them have $y_i = +1$ and those three come last in the ranking. what is your average precision?
- Suppose you have N total number of samples, R of them have $y_i = +1$ and those R come last in the ranking. what is your average precision? You will (likely) not be able to write it in a simple term like $\sum_{i=1}^{R} i = \frac{R(R+1)}{2}$. Give an expression which depends on R, N and contains a sum.
- Now consider a random predictor: suppose you have N total samples, R of them have $y_i = +1$ and those are on evenly distributed (bcs every time you train the predictor, it learns nothing) in the following sense (to give a simplified calculation):

  indexing starts at 1 and the first $y_i = 1$ sample appears at $\frac{N}{R}$, the second sample at $2\frac{N}{R}$, the l-th sample at $l\frac{N}{R}$, etc. What is the average precision then ? You **will be able** to write it as a simple term.

  Hint: calculate the precision@k for every index $k = l\frac{N}{R}$. Then the Average precision.
- Now consider a slightly shifted random predictor: suppose you have N total samples, R of them have $y_i = +1$ and those are on evenly distributed (bcs every time you train the predictor, it learns nothing) in the following sense (to give a simplified calculation):

  indexing starts at 1 and the first $y_i = 1$ sample appears at 1, the second sample at $1 + \frac{N}{R}$, the l-th sample at $1 + (l-1)\frac{N}{R}$, etc.

What is the average precision then ? Here you will not be able to write it in a simple term again. Will the average precision for this case be lower or higher than in the case before?

- consider a linear classifier $s(x) = wx + b$ with trainable parameters w and b. Consider (a) accuracy by 0-1-loss and (b) mean average precision. For (a) and (b) which of the trainable parameters have an impact on the result and why?