Task 1: Precision and Recall (theoretical)

Two search engines A and B perform a search on the same collection. Each engine returns the top 30 documents for a single query in ranked order by their relevance. The following table provides the ranking and denotes with a '+' if the document is relevant and with an empty cell if it is not relevant. The collection contains a total of 12 relevant documents for this single query.

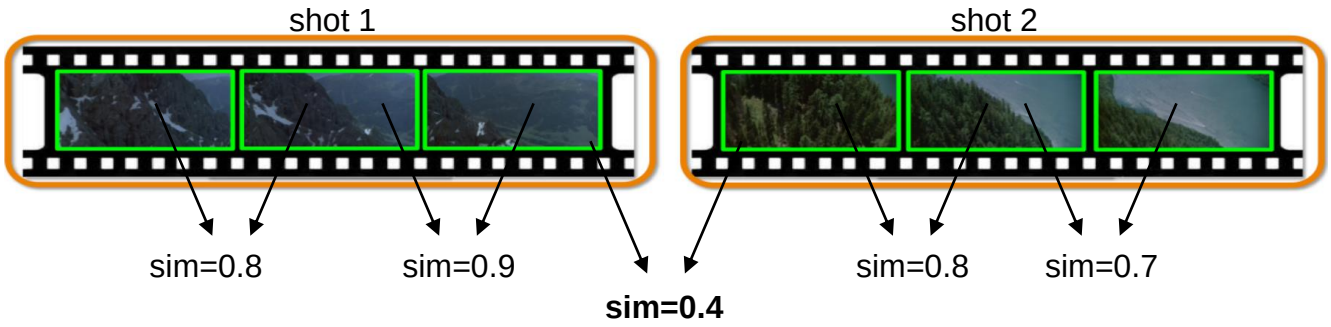
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
A	+	+		+				+				+								+										+
B		+	+				+			+	+		+				+			+			+		+				

a) Draw the precision-recall graph for both engines.

b) Which engine performs better and why?

Task 2: Performance of Video Shot Detection (practical)

We have written a software that segments a video into shots, i.e., a sequence of frames that belong together. The software computes similarity values between subsequent frames and assumes the end of a shot if the similarity value is below a threshold; otherwise it considers both frames belonging to the same shot:



In the example above, there are two shots each with 3 frames, and we are given the similarity values between subsequent frames. As we see, the similarity value of the two frames at boundary of the shots is much smaller than the ones between frames within a shot. If we use a threshold, for example, of 0.5, we can safely detect the shot segmentation.

We are now looking at a more realistic example. You find two text files on the courses homepage for two different methods of our shot detector. Each file contains the similarity values between two subsequent frames together with labels “shot” or “noshot” to indicate whether there is a shot boundary between the frames or not (the labels are given to train our system). This exercise is best done in Excel, but you can use any other means to compute the results.

a) Compute the true/false positives/negatives for the example data ordered by decreasing thresholds on the similarity values.

b) Plot the ROC curves and identify an optimal threshold for the shot detection. What is your measure of optimality? How well do both methods perform with their optimal threshold?

c) Compute the area under the ROC curve and determine which method works better. Is this the same ordering of the engines as for sub task b)?