

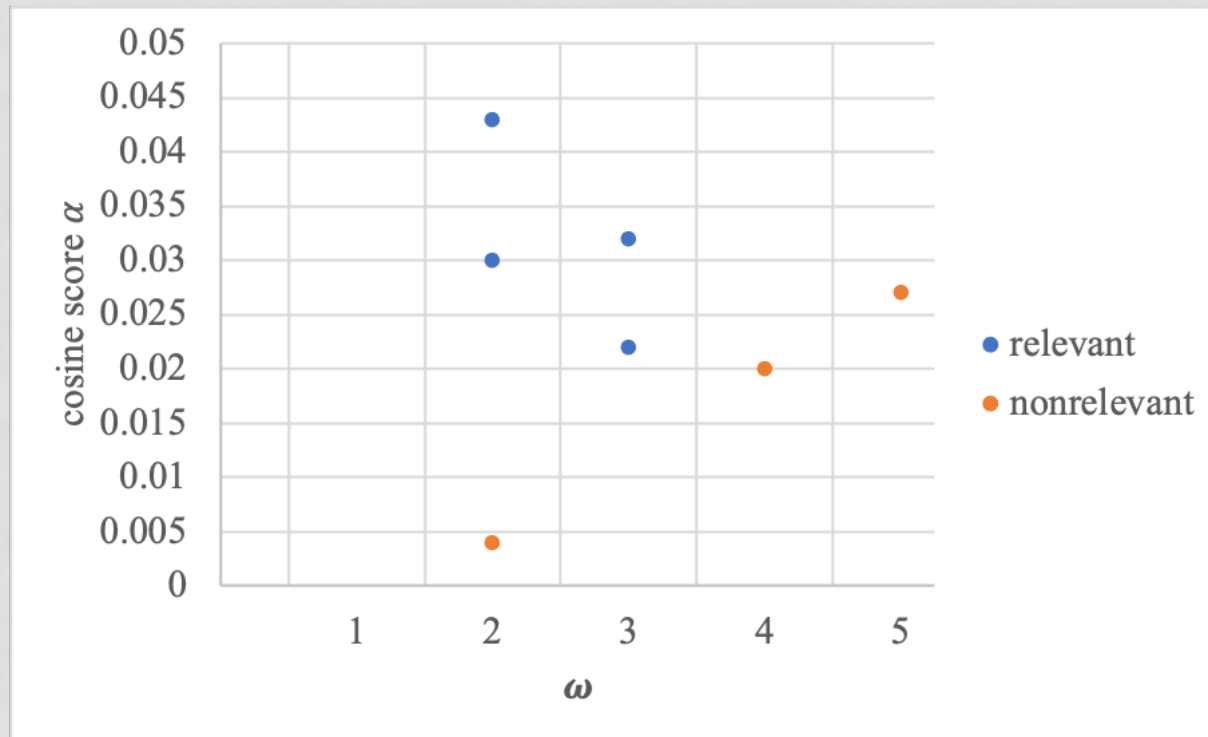
# INFORMATION RETRIEVAL

HOMEWORK EXERCISES L09. LEARNING TO RANK

SUZAN VERBERNE 2022

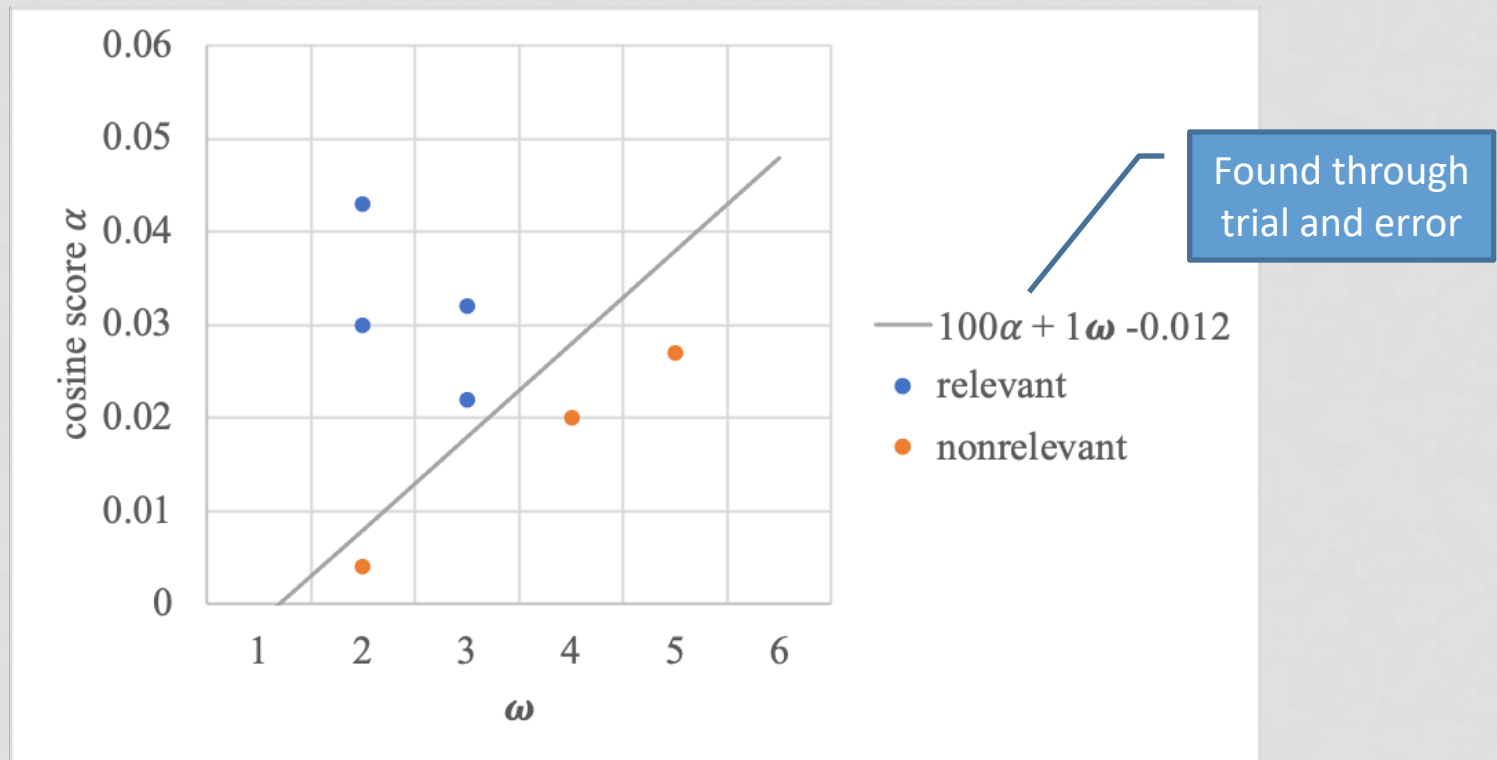
# EXERCISE 1 - SOLUTION

- IIR exercise 15.7: Plot the first 7 rows of Table 15.3 in the  $\alpha$ - $\omega$  plane to produce a figure like that in Figure 15.7.



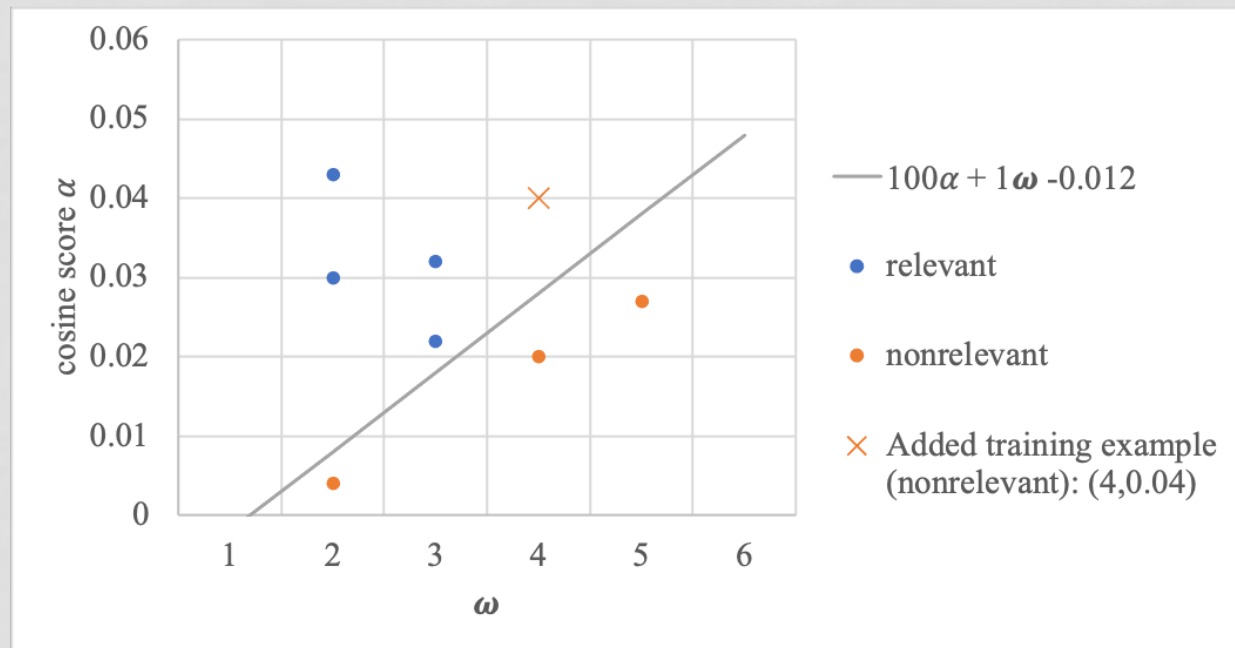
# EXERCISE 2 - SOLUTION

- IIR exercise 15.8: Write down the equation of a line in the  $\alpha$ - $\omega$  plane separating the Rs from the Ns.



# EXERCISE 3 - SOLUTION

IIR exercise 15.9: Give a training example (consisting of values for  $\alpha$ ,  $\omega$  and the relevance judgment) that when added to the training set makes it impossible to separate the R's from the N's using a line in the  $\alpha$ - $\omega$  plane.



# EXERCISE 4 - SOLUTION

a. Pointwise loss (squared regression loss)

	True: $rel_q(d)$	Model A output: $score_A(q, d)$	$\mathcal{L}_{squared,A}$	Model B output: $score_B(q, d)$	$\mathcal{L}_{squared,B}$
d1	1	0.2	0.64	0.6	0.16
d2	1	0.1	0.81	0.8	0.04
d3	0	0.1	0.01	0.2	0.04
d4	0	0.1	0.01	0.4	0.16
		Avg loss	<b>0.37</b>		0.10

# EXERCISE 4 - SOLUTION

b. Pairwise loss (hinge loss)

	True: $rel_q(d)$	Model A output: $score_A(q, d)$	$\mathcal{L}_{squared,A}$	Model B output: $score_B(q, d)$	$\mathcal{L}_{squared,B}$
			0.9		0.6
			0.9		0.8
			1.0		0.4
			1.0		0.6
		Sum loss	3.8		2.4

# EXERCISE 4 - SOLUTION

2. If you rank the documents based on the system scores, which ranking is better
  - a. according to the pointwise loss function
    - B is better (loss is lower)
  - b. according to the pairwise loss function
    - B is better (loss is lower)
  - c. according to you based on the true relevance values
    - Probably B. A has three documents with the same score (0.1), one of which is relevant. With only knowing these scores, the ranking of these three documents will be random and the relevant one might end up at the bottom. In B, the relevant documents have higher scores than the nonrelevant documents.