

Homework Assignment 4

Exercise 4.1

In language models one can estimate a model's probability parameters by the relative term frequencies as they appear in the respective document. What is the problem with this approach and how can it be fixed?

Exercise 4.2

Using an example of your own, explain the difference between relevance and pertinence.

Exercise 4.3

In the lecture (slide 52; "endangered species") we showed an example information need as used in TREC. Think of an information need of your own and describe it using the same structure (title, description, narrative).

Exercise 4.4

What is the pooling method?

Exercise 4.5

What is more important for an information retrieval system: precision or recall?