1. Consider a search need for which there are 5 relevant documents in the collection. Two information retrieval systems returned 10 documents that were judged according to their relevance as follows:

   a. What is the P@10 of each system?

   b. What is the Recall of each system?

   c. What is the Average Precision (AP) of each system?

   d. What is the DCG@3 of each system?

   e. Considering the above results, what does each metric allows you to conclude?

Rank	S1	S2
1	2	1
2	1	2
3	1	0
4	0	0
5	1	0
6	0	1
7	0	0
8	0	2
9	0	0
10	0	0

2. Consider the TF-IDF term weighting.

   a. When is idf greater (or smaller) then 1? (consider log base 2)

   b. What is the IDF of a term that occurs in every document? Compare this with the use of stop word lists.

3. Consider the table of term frequencies for 3 documents. The collection contains: 50,000 documents, 200,000 unique terms and 500,000 total number of terms.

	docFrequency	Doc1	Doc2	Doc3	
car	18.165	13	4	22	
auto	6.723	2	0	33	
insurance	2.000	0	39	33	
best	435	7	0	17	
		300	789	1000	Length

   a. Compute the LMD probabilities for the above terms for all documents.

   b. For the query “best insurance” which document is more relevant?

   c. For a given search, the original query was the term “best” and Doc2 was selected as a positive example. If you apply the Rochio algorithm with the first document only, what will be the new query?

4. Rank fusion methods combine ranks in different ways. Consider the table below with different ranks.

Rank 1 (id/score)	Rank 2 (id/score)	Rank 3 (id/score)
D3 / 0.5	D3 / 1.6	D1 / 15.0
D4 / 0.2	D8 / 0.8	D2 / 10.0
D2 / 5.0	D2 / 0.8	D8 / 8.0
D5 / 0.18	D1 / 0.5	D3 / 0.19

a. Compute the fused rank for the above lists with the RRF method.

b. Compute the fused ranks for the above lists with the wCombMNZ method.

c. Justify the values you assign to each document that is not present in one of the lists

5. Suppose that your boss asks you to create a collection to develop a new corporate search engine for your company.
The company is currently running a 5 years old ElasticSearch with LMD and whishes to replace it with a BERT enabled search engine.
The test collection needs to be useful for the next 2-3 years.

a. Describe how you would build or acquire the different collection components (e.g., test, train, queries, relevant docs per query), and how much data is required for each component.

b. Detail the process of selecting queries and acquiring the corresponding relevance judgments. Your answer needs to be practical, i.e., no magic, and your budget isn't infinite.

6. Justify if the following statements are true or false:

a. Words that co-occur in a given sentence will always be near in the embedding space.

b. The average distance between two words in the corpus sentences is used to select negative examples when learning word embeddings.

09:25

6. Justify if the following statements are true or false:

   a. In word embeddings, words are represented as vectors, where a word frequency is not relevant.

   b. In word embeddings, words are represented as vectors where the neighborhood of a word vector *i*, is linked to the words tha co-occurring in the same sentences as word *i*.

14:

7. Consider the Conversational Search agent of the course project.

   a. Detail the different components of the implemented pipeline.

   b. Describe the required steps to train your re-ranking classifier.

   c. Implement the code to print the contents of all the documents of relevance equal to 1 in the test set.

13:18

7. Consider the Conversational Search agent of the course project.

   a. Detail the different components of the implemented pipeline.

   b. Which queries were used for the classifiers that you implemented for the state tracking methods?

   c. Implement the code to print the contents of all the documents of relevance equal to 2 in the test set.