Recuperação de Informação / Information Retrieval
2018/2019 MIECT/MEI, DETI, UA

Assignment 3
Submission deadline: **3 December 2018**

Calculate performance metrics for your weighted retrieval method using the Cranfield corpus.

1. Index the corpus ☺

2. Process the queries (file 'cranfield.queries.txt') and retrieve the sorted results for each query.

3. Using the relevance scores (file 'cranfield.query.relevance.txt') provided, calculate the following evaluation and efficiency metrics:

   a. Precision
   b. Recall
   c. F-measure
   d. Mean Average Precision
   e. Mean Precision at Rank 10
   f. Normalized Discounted Cumulative Gain (NDCG)

   i. Query throughput
   ii. Median query latency

Suggestions:

   – Write **modular** code

   – Favour **efficient** data structures

   – Add **comments** to your code

   – Follow the **submission instructions**

Submission instructions:

- To manage your project please use **<u>Maven</u>**

- At each submission, include a small **<u>Report</u>** including:

  o Your project's **<u>class diagram</u>**

  o Complete instructions on how to run your code, including any parameters that need to be changed

  o A list of any external libraries that are needed to run the code

- Make sure you **include your name and student number** in the code and in the report.

- Make sure all your programs compile and run correctly.

- Submit your assignment by the due date using Moodle.