Report Information Retrieval

Assignment 2 [Vinit Raj, 19CS10065]

Part 2A
Helped in Debugging the final code that was written by other members:
Problem with NaN values and config of tf-idf calculations.

Implementations:
Part 2B(Evaluation)
Preprocessing:
    After taking in the input from the files, the ground truth list is modified into a list of dictionaries [35 x (dict {cord-id → (judgment_score, last_iteration)})]
This allows easier access to the scores while traversing the ranked output of our IR system.

Mean Average Precision Calculation:
    In the function for MAP calculation, first, we pass the ranked list of each query, the ground truth and the K value.
Then simply iterate through the ranked list accumulating the precision only at relevant docs, this is done by maintaining a running count of relevant docs.
Once done we average it and get the average precision for that particular query.
We accumulate this across the queries and taking mean again would give us the Mean Average Precision@K.