Information Retrieval HW2

1. System model: Vector space model
   mean average precision: mAP = 0.49692164
2. System model: BM25
   mean average precision: mAP = 0.5045349
3. Feedback
   I chose to build up vector space model to do the document ranking at first. However, the map never reaches 0.5 no matter how I adjust the function of the tf-idf. Among those experiments, I found that in this case setting tf as simple frequency can generate better result as showed above. I supposed the reason may be the small collection of document and no process to remove the effect of stop words. Because the collection provided are namely number, it is impossible for us to eliminate meaningless and common words. Though idf can decrease the influence of those words on the total ranking, yet I guess there are a huge amount of those words existing among the query and documents that it lowers the effectiveness of adjustment from idf.

   Given the failure of vector space model, I then turned to BM25 to expect a better result. The outcome of the map is over 0.5 slightly. Although it seems the gap between these two medel is minor, it still indicates BM25 is more suitable for this case than the vector space model.

   From this experience, I figured out that the model choosing is important for different document collection. Some words don't show up as much as those stop words, but they are not the critical one either to decide the relevance between the document and the query. If the vector space model is used in this context, worse ranking result is expected to happen. If BM25 is used instead, the function may adjust the weight to make the critical terms more standing out than others and ensures the ranking more closely reveal the relevance to the query.