

MOTEURS DE RECHERCHE

M2 IAD

TME 2 - Modèles de recherche et évaluation

L'objectif de ce TME est de programmer et d'évaluer l'ensemble des modèles de recherche vus en cours, sur les deux collections de documents utilisées lors du premier TME.

Plus précisément, l'objectif du TME est le suivant :

- adapter votre procédure d'indexation pour garder en mémoire les tf des termes dans les documents,
- avoir des procédures de calcul des idf au moment de l'indexation,
- programmer les différents modèles de recherche (somme des $tf.idf$, modèle vectoriel, BM25 et modèles de langage),
- programmer les différentes mesures d'évaluation ($prec@k$ pour différentes valeurs du rang k , MAP),
- évaluer chacun de ces modèles sur les bases CACM et CISI en utilisant les requêtes et les jugements de pertinence qui sont fournis,
- trouver les paramètres optimaux pour BM25 et le modèle de langage,
- tracer les courbes de rappel/précision pour chacun des modèles, et en déduire les modèles de recherche les plus performants sur les deux collections.