Introduction générale

La recherche d’information est un domaine historiquement lié aux sciences de l’information et à la bibliothéconomie. Elle traite l’information dans la manière de l’organiser et de la façon de la sélectionner, elle peut être définie comme une activité qui dans le but de répondre à une question vise à localiser et à traiter une ou plusieurs informations au sein d’un environnement documentaire complexe.

Le traitement de cet environnement ne peut être effectué manuellement et donc l’objectif de la recherche d’information est d’extraire les informations pertinentes vis-à-vis d’une requête pour un utilisateur donné à travers l’utilisation d’un ensemble de programmes informatiques appelés systèmes de recherche d’information.

Un SRI peut guider l’utilisateur vers une bonne formulation de ses besoins. La solution proposée dans le but de réduire la distance entre la pertinence système et la pertinence utilisateur est la combinaison du score initial de la probabilité de pertinence document/requête avec le score à priori du document

Dans ce mémoire on s’intéresse particulièrement à l’extension d’un modèle de recherche en utilisant le score à priori de document, pour cela il faudra calculer les caractéristiques pris en compte pour le calcul de pertinence à priori de documents, puis combiner ce score avec le score de vraisemblance Document/Requête, pour enfin présenter les résultats obtenus par notre approche.

Pour atteindre cet objectif, nous avons structuré ce mémoire en trois chapitres :

Le premier chapitre intitulé « La recherche d’information » présente globalement la discipline de la recherche d’information.

Le deuxième chapitre intitulé « La pertinence à priori de documents » présente les caractéristiques prises en compte pour effectuer un classement à priori de documents ainsi que la méthode de combinaison de scores (à priori de document et score de vraisemblance Document/Requête).

Le dernier chapitre est consacré à la présentation de notre approche d’extension d’un modèle de recherche d’information en utilisant le score à priori de documents, ainsi que quelques résultats expérimentaux obtenus.