

Ecole SITI
Département
« IMATH »
Ingénierie
Mathématique

Chaire de
Statistique appliquée

Séminaire de Statistique appliquée

mercredi 21 mars 2012 de 11H à 12H

en salle 17.2.15

(292 rue Saint-Martin– Paris 3^{ème})

**« Text mining et système de recommandation pour les
offres d'emploi diffusées sur le web »**

présenté par **Julie Séguéla**

Doctorante CNAM Département IMATH

Résumé

L'expansion du média Internet pour le recrutement a entraîné ces dernières années la multiplication des canaux dédiés à la diffusion des offres d'emploi. Évaluer et comparer les performances des différents canaux de recrutement est donc devenu un besoin pour les entreprises. Dans ce contexte, notre objectif est de fournir au recruteur une indication de la performance attendue sur les sites d'emploi pour un poste à pourvoir donné. Après avoir identifié les facteurs explicatifs potentiels de la performance d'une campagne de recrutement, nous appliquons aux annonces des techniques de fouille de textes afin de les structurer et d'en extraire de l'information pertinente pour enrichir leur description au sein d'un modèle explicatif.

Nous proposons dans un second temps un algorithme prédictif de la performance des offres d'emploi, basé sur un système hybride de recommandation, adapté à la problématique de démarrage à froid. Ce système, basé sur une mesure de similarité supervisée, montre des résultats supérieurs à ceux obtenus avec des approches classiques de modélisation multivariée. Nos expérimentations sont menées sur un jeu de données réelles, issues d'une base de données d'annonces publiées sur des sites d'emploi.

Contact « Séminaire de Statistique appliquée » :

Mr Pierre-Louis Gonzalez
Cnam - Département IMATH
E.mail : pierre-louis.gonzalez@cnam.fr
Site web : <http://www.maths.cnam.fr/> -
Rubrique : « la Recherche ; Séminaires de Statistique »