

Sujets Master 2 Professionnel

Fonctionnement

Les étudiants de Master 2 Professionnel doivent se répartir en 2 groupes. Chaque groupe doit choisir l'un des sujets ci-dessous (il n'est pas nécessaire que les sujets choisis soient différents). Etant donné les description succincte du sujet qu'il aura choisi, chaque groupe devra :

- Préciser le contexte du sujet, l'objectif visé.
- Décrire un algorithme pour aborder le problème posé par le sujet.
- Récupérer un petit échantillon de données sur lequel exécuter l'algorithme retenu.
- Evaluer la qualité de la méthode utilisée, en en commentant les limites et les aspects positifs.

Chaque groupe devra rédiger un rapport de 6 pages maximum où ces différents points devront être abordés (des modèles de rapports peuvent se trouver sur la page : <http://www.stanford.edu/class/cs229/projects2006.html>).

Le travail de chaque groupe fera l'objet d'une présentation de 50 minutes (avec démonstration, si possible) mercredi 24 février 2010 au matin. Lors de la présentation, le groupe qui ne présentera jouera le rôle de « contradicteur » : il posera des questions sur le travail effectué, sur les limites qu'il semblera avoir détectées sur l'approche choisie et proposera éventuellement quelques pistes qui auraient pu améliorer le projet présenté. Ce rôle de contradicteur compte dans l'évaluation, pour près de la moitié de la note finale.

Un encadrement d'1h30 (obligatoire) par semaine le mercredi sera assuré tout au long de la préparation de ce mini-projet. Le planning précis est en cours de réalisation.

Netflix

Le projet Netflix concerne la mise au point d'algorithmes pour la recommandation de films. Parmi ces algorithmes, ceux reposant sur la factorisation de matrices sont les plus performants. Le problème de la recommandation sera précisément vu sous cet angle.

Catégorisation de dépêches de presse

Le problème de la catégorisation de dépêches de presse est un problème de fouille de textes qui peut être abordé à l'aide de méthodes d'apprentissage automatique. C'est le cadre dans lequel se place ce sujet.

Are you hot or not ?

Le site <http://www.hotornot.com/> met à disposition des millions d'images de personnes. La « beauté » (l'usage des guillemets provient de ce que la beauté est une notion toute subjective) de ces personnes est évaluée par des milliers d'internautes. La question que l'on se pose ici est celle de la « beauté » computationnelle, à savoir celle de savoir si un ordinateur est capable d'identifier les personnes jugées comme belles. (Note : on ne travaillera que sur les visages.)