Plan de cours: IFT 6266
Algorithmes d'apprentissage
www.iro.umontreal.ca/~pift6266
Professeur: Yoshua Bengio

La recherche sur les algorithmes d'apprentissage statistique a mené à d'importantes découvertes au cours des deux dernières décennies, qui ont changé la manière de concevoir le problème de rendre les ordinateurs plus intelligents. Ces algorithmes ont pour objectif de permettre à la machine d'apprendre à partir d'exemples. Ils cherchent à découvrir la structure cachée d'un ensemble de signaux à partir des valeurs observées. Ils sont utiles pour la prédiction, la prise de décision, et l'analyse de données. Ils ont mené à de nouveaux champs commerciaux, comme le "data-mining", l'analyse statistique de données textuelles pour le web, ou la gestion quantitative de porfefeuilles financiers. Ce cours est complémentaire au cours IFT6390 (mais si les deux cours sont suivis il est recommandé de commencer par IFT6390). Les sujets principaux qui seront traités sont les suivants:

- Dans l'apprentissage supervisé, on présente à l'ordinateur des exemples sous la forme de paires (entrée, sortie désirée). Les deux formes principales d'apprentissage supervisé sont la classification supervisée (la sortie désirée est une catégorie discrète) et la régression (la sortie désirée est un ou plusieurs nombres réels). On parlera plus ici de régression, qui dans le cas classique consiste en l'estimation de l'espérance conditionnelle, et de son extension à des fonctions de coût plus générales, qui sont par exemple utiles en finance.

- Le but fondamental de l'apprentissage est de bien généraliser à de nouveaux cas. Formalisation mathématique du problème. Capacité d'une classe de fonctions. Fonctions de coût. Régularisation. Bornes asymptotiques et non-asymptotiques sur l'erreur de généralisation.

- Dans l'application pratique de ces algorithmes il est très important de maîtriser les méthodes statistiques qui permettent d'évaluer la qualité d'un prédicteur obtenu à partir de données, afin de savoir si il vaut la peine de l'employer et si il est vraiment supérieur à des méthodes de référence. Certaines questions ouvertes concernant ce problème seront abordées.

- Modèles génératifs probabilistes et apprentissage non-supervisé: principe du maximum de vraisemblance, principes d'inférence Bayesienne, mélange de Gaussiennes, malédiction de la dimensionalité, algorithme EM, réduction de la dimensionalité par modélisation de variétés non-linéaires, modèles probabilistes graphiques.

- Algorithmes d'apprentissage non-paramétriques classiques et modernes (incluant les réseaux de neurones et les modèles à noyau: SVMs et processus Gaussiens).
- Apprentissage de données temporelles: modèles de Markov cachés, filtre de Kalman, réseaux de neurones récurrents, données non-indépendentes et non-stationnaires. Exemples: données financières, langage, reconnaissance de la parole, séquences bio-informatiques.

Modalités d'évaluation

L'évaluation de l'étudiant se fera au moyen de deux examens (un intra et un final), et de devoirs théoriques et pratiques. Les devoirs seront données environ à chaque deux semaines, sauf le travail final qui sera plus substantiel et impliquera la comparaison et l'évaluation d'algorithmes d'apprentissage. Les examens compteront pour 50% des points (25% chacun pour l'intra et le final) et les devoirs pour le reste des points. Le travail final comptera pour 20% des points.

Support de cours

Le support principal du cours sera le site web `www.iro.umontreal.ca/~pift6266`. On y trouvera des notes de cours, des pointeurs sur des articles disponibles électroniquement sur le web, les énoncés des devoirs ainsi que toute autre information importante.
Le livre de référence de base pour le cours est:

- Hastie, Tibshirani, Friedman, *The Elements of Statistical Learning*, Springer.

Il a été commandé à la librairie de l'Université. D'autres ouvrages seront suggérés sur le site de cours et mis en réserve à la bibliothèque.

Horaire

Première rencontre : mercredi 3 septembre 2003.

Jour	**Heure**	**Lieu**	**Type**
Lundi	15h30-17h30	Z-337	Cours magistral
Mercredi	15h30-17h30	1177	Cours magistral