

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. Un agent est intelligent parce qu'il a des connaissances opérationnelles (pas nécessairement sous une forme verbale explicite) qui lui permettent d'effectuer certaines tâches ou de répondre à certaines questions sur un certain domaine. Les algorithmes d'apprentissage ont pour objectif de permettre à la machine d'acquérir des connaissances opérationnelles à partir d'exemples. Ils cherchent à découvrir la structure cachée d'un ensemble de signaux à partir des valeurs observées, les facteurs inhérents qui expliquent les variations et les dépendences 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 (à la base des engins de recherche web modernes) ou la gestion quantitative de portefeuilles financiers. Ce cours est complémentaire au cours IFT6390 (mais si les deux cours sont suivis il est recommandé de commencer par IFT6390). Les étudiants moins à l'aise en mathématiques devraient prendre IFT6390 (à la place ou avant IFT6266). 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. Comportement asymptotique et non-asymptotique de l'erreur de généralisation quand le nombre d'exemples augmente.
- 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.
- Modèles génératifs probabilistes et apprentissage non-supervisé: principe du maximum de vraisemblance, principes d'inférence Bayésienne, 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, processus Gaussiens, apprentissage non-paramétrique de variétés).
- Apprentissage de données temporelles: modèles de Markov cachés et autres réseaux Bayesiens temporels, réseaux de neurones récurrents, données non-indépendantes et non-stationnaires. Exemples: données financières, données textuelles, reconnaissance de la parole, séquences bio-informatiques, vidéo.

Prérequis

Il y n'y a pas de prérequis formels, mais on recommande une certaine capacité à programmer rapidement des prototypes dans le langage de votre choix, ainsi qu'une certaine aisance en mathématiques (particulièrement probabilités, statistiques, algèbre linéaire, et dans une moindre mesure optimisation). Et un esprit ouvert et curieux.

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 donnés 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

Les deux supports principaux pour le cours seront d'un part le site web

www.iro.umontreal.ca/~pift6266

et d'autre part le livre de Bishop ci-bas. On trouvera sur le site web 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 pour le cours est le suivant:

- Bishop, C., *Pattern Recognition and Machine Learning*, Springer, 2006.

Horaire

Première rencontre :
Mercredi 3 septembre 2008,
10h30, au Z-209.

Jour	Heure	Lieu	Type
Lundi	10h30-12h30	Z-209	Cours magistral
Mercredi	10h30-12h30	Z-209	Cours magistral