

IFT 6266

Algorithmes d'apprentissage

Site WEB du cours:

<http://www.iro.umontreal.ca/~bengioy/ift6266>

Yoshua Bengio

Bureau : Pavillon Aisenstadt #3339, e-mail: bengioy@iro.umontreal.ca, #tel: 343-6804

1 Objectifs

Ce cours est une introduction à un champ de recherche en Intelligence Artificielle qui a connu une forte croissance dans les dernières années. L'objectif central des algorithmes étudiés est l'apprentissage automatique: comment apprendre une tâche avec un nombre fini d'exemples. Nous considérons ici trois classes de problèmes: reconnaissance de formes, reconnaissance ou prédiction de séquences et de séries chronologiques, estimation de densité de probabilité. Une aisance en maths et des connaissances de base en statistiques et probabilités seront fort utiles à l'étudiant.

Le cours commencera par un survol des principaux éléments de la théorie de l'apprentissage, qui concerne la question: qu'est-ce qui fait qu'un algorithme appliqué à un ensemble d'exemples fonctionnera bien, donc généralisera à de nouveaux exemples? Pour cela nous définirons formellement la notion d'apprentissage comme choix d'une fonction parmi un ensemble de fonctions, celui de généralisation comme d'une performance espérée, et nous introduirons la notion de capacité qui permet de borner l'erreur de généralisation.

Trois classes principales d'algorithmes seront ensuite étudiés: des algorithmes classiques de reconnaissance des formes et de régression à partir de modèles statistiques, des algorithmes de *réseaux de neurones artificiels* (inspirés d'analogies avec le traitement de l'information dans le cerveau), et des algorithmes tels que les modèles de Markov cachés, qui cherchent à utiliser la structure temporelle des données. Nous parlerons aussi d'algorithmes de plus haut niveau qui permettent d'améliorer les performances en combinant plusieurs modèles.

Une des raisons de l'engouement récent pour les approches introduites dans ce cours est leur succès initial rapide pour diverses applications, en exploitation de grandes bases de données (data-mining), en reconnaissance d'image, de caractères, de la parole, en modélisation du langage, en prédiction de séries temporelles, en contrôle de processus industriels, et dans la prise de décision en général. Nous illustrons certains de ces algorithmes avec des exemples tirés de ces applications. De plus, un élément important du travail de l'étudiant sera un projet pratique sur l'application d'un algorithme d'apprentissage. L'étudiant sera aussi évalué sur la base de devoirs théoriques ainsi que d'un examen à la mi-session.

2 Support de cours

Les notes de cours sont ou seront sur le site Web du cours. Une référence de base est fortement suggérée aux étudiants:

- Bishop, C., *Neural networks for pattern recognition*, Oxford University Press, 1995. (disponible pour environ 70\$ à la librairie de l'UdeM).

Les ouvrages suivants pourront aussi être utiles:

1. Rice, J.A., *Mathematical Statistics and Data Analysis*, (en réserve à la bibliothèque).
2. Haykin, S., *Neural Networks: a Comprehensive Foundation*, (voir à la librairie de l'UdeM).
3. Bengio, Y., *Neural Networks for Speech and Sequence Recognition*, Thomson Computer Press, 1996.
4. Parallel Distributed Processing: Explorations in the Microstructure of Cognition - Volume 1: Foundations, D. E. Rumelhart and J. L. McClelland, eds., Bradford Book, MIT Press, 1986.
5. Krogh and Palmer, *Introduction to the Theory of Neural Computation*, Addison-Wesley, 1991.
6. Duda and Hart, *Pattern Classification and Scene Analysis*, Wiley-Interscience, 1973.

3 Horaire

Première rencontre :
28 août 2000.

Jour	Heure	Lieu	Type
Lundi	9h30-11h30	1207	Cours magistral
Mardi	10h30-12h30	1411	Cours magistral

4 Devoirs et Travaux pratiques

	Thème	Donné	Rendu
TP #1	Capacité et généralisation (théorique)	05/09	18/09
TP #2	Descente de gradient (théorique)	19/09	02/10
TP #3	Projet (programmation)	03/10	5/12

Les dates sont approximatives.

5 Évaluation

Mode d'évaluation	Épreuve	Valeur
Examen	Intra	30
Travail pratique	TP1	10
Travail pratique	TP2	10
Projet	TP3	50

6 Plan du Cours

1. Introduction et motivation

- apprentissage et généralisation,
- reconnaissance des formes et séries chronologiques.

2. Apprentissage automatique

- généralisation, capacité.
- bornes sur l'erreur,
- dilemme biais/variance, intégration de connaissances a-priori,
- réseaux Bayesiens,
- inférence sur l'erreur de généralisation.

3. Algorithmes de reconnaissance des formes:

- classifieurs linéaires et polynomiaux,
- SVM ("support vector machines"),
- analyse statistique non-paramétrique classique,
- arbres de décision.

4. Réseaux de neurones:

- réseaux multicouche, rétropropagation de l'erreur,
- connaissances a-priori $\rightarrow$ contraintes,
- partage de poids et réseaux à convolution: TDNN, LeNet,
- optimisation numérique non-quadratique.

5. Modèles de la structure temporelle

- modèles de Markov cachés,
- réseaux de neurones récurrents,
- problème de l'assignation du crédit à travers le temps,
- problème de la non-stationarité.

6. Mixtures, ensembles, et combinaison de modèles:

- machines hybrides, systèmes multimodulaires,
- mixtures d'experts,
- bootstrap et intervalles de confiance,
- ensembles de modèles, "bagging",
- l'algorithme "Adaboost".