

Plan de cours: IFT6266 H2016

Algorithmes d'apprentissage (de représentations)

<https://ift6266h16.wordpress.com>

Professeur: Yoshua Bengio (<http://www.iro.umontreal.ca/bengioy>)

Assistant d'enseignement: Bart Merriënboer (bart.vanmerrienboer@gmail.com)

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 configurations de valeurs observées, donc à découvrir 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. La recherche en intelligence artificielle a elle aussi été fortement influencée par l'apprentissage automatique, et c'est ce type d'applications (vision, langage, musique, parole, etc.) qui sera visé dans la partie pratique du cours.

Ce cours est un deuxième cours sur les algorithmes d'apprentissage, par exemple pour faire suite au cours IFT3395 ou IFT6390. Les étudiants qui n'ont jamais pris de cours en apprentissage automatique devraient donc plutôt commencer par IFT6390 (ou l'équivalent) avant de prendre IFT6266.

Les sujets couverts en IFT6266 sont des sujets plus avancés, et liés à certaines découvertes faites au cours des dernières années (dont certaines au DIRO), en particulier dans le domaine de l'apprentissage profond (*Deep Learning*) et de l'apprentissage de représentations (qui ont mené à la création d'une nouvelle conférence, ICLR, International Conference on Learning Representations).

- Rappels sur les réseaux de neurones
 - Perceptrons, régression linéaire et logistique
 - Rétropropagation et optimisation à base de gradient
- Réseaux de neurones avancés:
 - Réseaux à convolution (données séquentielles et spatiales).
 - Réseaux récurrents (en particulier le LSTM).
 - Auto-encodeurs et variantes.
 - Hyper-paramètres et trucs pour l'entraînement des réseaux de neurones.
 - Difficulté d'entraîner des architectures profondes.
- Modèles probabilistes graphiques:
 - Rappel sur les modèles dirigés (HMMs et mélanges).
 - Modèles non-dirigés: champs Markoviens et machines de Boltzmann.
 - Méthodes d'inférence et méthodes d'échantillonnage, en particulier pour les machines de Boltzmann restreintes.
- Réseaux profonds:

Plan de cours: IFT6266 H2016

Algorithmes d'apprentissage (de représentations)

<https://ift6266h16.wordpress.com>

Professeur: Yoshua Bengio (<http://www.iro.umontreal.ca/bengioy>)

Assistant d'enseignement: Bart Merriënboer (bart.vanmerrienboer@gmail.com)

- Apprentissage non-supervisé de représentations.
- Pré-entraînement non-supervisé, Deep Belief Networks, machine de Boltzmann profonde.
- Variations sur les auto-encodeurs pour l'apprentissage de représentations et les modèles génératifs.
- Variantes architecturales modernes pour les images et les séquences.

Modalités pédagogiques et évaluation: Le cours se basera sur plusieurs sources de matériel pédagogique. Une grande partie viendra du cours de Geoffrey Hinton sur Coursera, intitulé Neural Networks for Machine Learning et de celui de Hugo Larochelle à l'Université de Sherbrooke, IFT725, Réseaux neuronaux. Les vidéos, articles et quiz seront étudiés par les étudiants avant de venir en classe, puis discutés en classe. Chaque étudiant aura la responsabilité de préparer des questions avant chaque cours (affichée sur le site du cours). Les discussions en classe sur ces questions permettront de clarifier les concepts, et de les compléter. Pour chaque question discutée, un étudiant sera chargé d'en transcrire la réponse sur le site du cours. En plus d'une série de notes de cours, plusieurs articles de survol du domaine serviront de base technique plus solide pour ces compléments. Des quiz et un examen final permettront d'évaluer les connaissances théoriques des étudiants sur ces sujets. Finalement, toute cette matière serait rapidement oubliée si elle n'était pas ancrée par une pratique régulière. Les étudiants devront expérimenter avec certains de ces algorithmes d'apprentissage de représentation, et ils compétitionneront entre eux sur des jeux de données choisis par le professeur, à travers l'interface Kaggle. Ils auront comme mission de maintenir un journal (visible en ligne par le professeur) documentant les détails de leurs expérimentations pour s'attaquer à ces tâches. Ils seront évalués à la fois sur leur succès dans cette compétition et sur la rigueur scientifique démontrés dans la démarche documentée par leur journal. Le grand prix sera de convertir les meilleurs documents et résultats en article de niveau suffisant pour être soumettre de manière crédible à une conférence majeure en apprentissage (comme NIPS, début juin). L'examen comptera pour 30% des points, les quiz, discussions et questions / réponses pour 30%, et le travail pratique pour 40%.

Horaire:

Première rencontre: Jeudi 7 janvier 2014, 9h30, au Z-305 du Pav. Claire-McNicoll.

Jour	Heure	Lieu
Lundi	14h30-16h30	Z-330 Pav. Claire-McNicoll
Jeudi	09h30-11h30	Z-305 Pav. Claire-McNicoll