Appel à Assistant d’enseignement : Cours/TDs dans le domaine de l’IA

Nous cherchons un doctorant ou postdoctorant, éventuellement un étudiant de M2, aux compétences avancées en apprentissage machine, pour un enseignement destiné à des étudiants en masters de sciences sociales. Les étudiants auront eu une initiation à Python et/ou R, à l’analyse de données et à l’apprentissage machine, mais il faut s’attendre à de fortes hétérogénéités de compétences. Le principe général du cours est exposé ci-dessous, les modalités d’organisation sont à discuter avec l’Assistant.

Période : second semestre 2020-2021, les lundis matins (jour et dates exactes à préciser)

Volume : 27h (HTD), incluant 18h de cours/tds (9 fois 2h ou 6 fois 3h), préparation, suivi des étudiants, contribution à la validation.

Lieu : Ecole Nationale des Chartes, 65 rue de Richelieu, 75002 Paris.

Contact : Jean-Pierre Nadal jpnadal@ehess.fr

--------------------------------------

IA pour les SHS : Datacamp/Data Challenge pour les étudiants des masters Humanités numériques (PSL) et Sciences économiques et sociales (EHESS-PSL).

Organisation : Jean-Pierre Nadal (CNRS & EHESS)

Comité de pilotage : Julie Josse (Ecole Polytechnique), Jean-Pierre Nadal (CAMS, CNRS & EHESS), Annick Vignes (CAMS,  Irstea)

Objectif : apprendre en faisant, en se confrontant à une compétition de type "Kaggle" (<https://www.kaggle.com/>). Ce type de compétition est devenu la norme dans le domaine de l’IA.  L’objectif du cours est de sensibiliser les étudiants à la science des données au sens large, avec une démarche assez classique où on part d’une problématique associée à un jeu de données et réaliser les différentes étapes d’une analyse en apprenant les méthodes, logiciels associés. On commencera donc par des étapes d’exploration, visualisation des données, puis des modélisations en utilisant les méthodes/algorithmes nécessaires pour répondre à la question (en insistant sur les compromis pouvoir prédictif/interprétabilité, la nécessité de toujours se comparer à des méthodes simples, etc). Enfin les étudiants communiqueront les résultats par exemple en réalisant des interfaces interactives.

Par ailleurs, nous nous appuierons également sur des ressources en lignes, de type MOOC, qui sont extrêmement bien réalisés. Ils peuvent aussi aider pour l’apprentissage des logiciels R, python, Tensorflow, etc., et permettre à chacun de se former avec des cours adaptés à son niveau.

L’objectif de cet enseignement est aussi d’utiliser les compétences des étudiants, de les encourager à chercher d’autres sources de données, d’autres informations de contexte qui leur permettront de suggérer les meilleures solutions.

Des jeux de données d’intérêt pour les SHS seront proposés aux étudiants.

Le cours aura lieu au second semestre les lundis (horaire exact à confirmer), à l'Ecole nationale des Chartes (65, rue de Richelieu, Paris 2e), salle informatique (salle équipée de PCs sous Ubuntu 16.04, avec R + Rstudio, Python 3 + PyCharm, oXygen, etc.).

Mots clefs :

statistique, machine learning, apprentissage supervisé, non supervisé, économétrie