Transparence des algorithmes pour la décision

ECTS : 3

Volume horaire : 24

Description du contenu de l'enseignement :

Les algorithmes jouent un rôle de plus en plus fondamental dans tous les aspects de notre vie. Une des principales craintes soulevées par les algorithmes est leur manque de transparence de leur fonctionnement, ce qui pourrait amener à l'avènement d'une « société boîte noire ». De plus, cette transparence est un prérequis pour pouvoir analyser et expliquer les décisions prises par les utilisateurs ou des entreprises.

Le but de ce cours est de discuter et de sensibiliser les étudiants aux questions soulevées par la transparence et la responsabilité des algorithmes telles que par exemple :
Quelles sont les situations où la transparence et l'explicabilité sont nécessaires et comment les implémenter en pratique ?
Comment réconcilier le besoin des entreprises de protéger leur propriété intellectuelle avec le besoin des citoyens de pouvoir comprendre la logique derrière les algorithmes qu'ils utilisent ou qui produisent un effet sur leurs vies ?
Comment améliorer la transparence, la responsabilité ainsi que l'équité des décisions prises par des systèmes algorithmiques sans pour autant compromettre leur utilité.

On s'intéressera ici à une catégorie particulière d'algorithmes, ceux qui sont utilisés pour l'aide à la décision ou dans des traitements qui ont des incidences sur les comportements individuels, qui ont donc un effet normatif. A titre d'exemples, on peut citer :
Les algorithmes de classement, qui établissent des priorités, des recommandations : on pense évidemment aux algorithmes de présentation des résultats des moteurs de recherche, mais aussi à ceux qui sont utilisés pour classer les candidats à un poste, etc.
Les algorithmes de catégorisation, de classification, de profilage comme ceux qu'on met en œuvre pour détecter des profils de potentiels terroristes, de fraudeurs, etc.

Ce cours abordera les points suivants :
Rappels de notions du langage Python
Transparence, explicabilité et responsabilité des algorithmes : définitions, liens, avantages et limites.
Le bien-fondé de la transparence des algorithmes pour l'informatique décisionnelle. Aspects juridiques.
Cas d'étude 1 : Transparence des algorithmes dans le cadre des systèmes de recommandation : le cas du filtrage collaboratif
Cas d'étude 2 : Transparence des algorithmes dans le cadre de la classification : classification ordonnée et arbre de décision.
Cas d'étude 3 : Transparence des algorithmes dans le cadre de l'élaboration des classements ou palmarès : prise en compte simultanée des préférences des utilisateurs et des données statistiques
Mise en œuvre en langage python, à travers un mini-projet, d'un algorithme transparent pour la résolution d'un problème, à partir de données réelles disponibles sur des plateformes comme https://www.data.gouv.fr ou https://www.kaggle.com/

Compétence à acquérir :

Rappels de notions du langage Python
Transparence, explicabilité et responsabilité des algorithmes : définitions, liens, avantages et limites.
Le bien-fondé de la transparence des algorithmes pour l'informatique décisionnelle. Aspects juridiques.