Traitement de textes et d'images avec logiciel R

ECTS : 2

Enseignant responsable : Marie BESSEC (https://dauphine.psl.eu/recherche/cvtheque/bessec-marie)

Langue du cours : Français

Volume horaire : 18

Description du contenu de l'enseignement :

Ce cours est une initiation au traitement de textes et d'images avec R et Python.

Dans la partie 1, nous apprendrons à extraire de l'information d'un recueil de textes (articles de presse, messages sur les réseaux sociaux, archives de presse, …). Après avoir nettoyé et numérisé le corpus, nous analyserons la fréquence des mots et leur association, la proximité et la tonalité des textes. Les techniques présentées nous permettront par exemple d'analyser l'optimisme des investisseurs à partir des messages postés sur les réseaux sociaux, de construire des indices de troubles sociaux ou de risque géopolitique à partir de l'analyse de la presse d'un pays ou d'identifier la tendance politique d'un texte.

Dans la partie 2, nous verrons les bases du Deep Learning appliquées à la vision informatique. Nous présenterons les réseaux de neurones et plus particulièrement les réseaux de neurones à convolutions, indispensables à l'analyse d'image. Nous mettrons en place en cours plusieurs types d'architecture de réseaux différents pour évaluer leurs performances respectives. Cette seconde partie vous donnera les bases pour mettre en place des algorithmes permettant de détecter des objets sur de l'imagerie satellite (voitures, maisons, piscines etc.) qui sont par exemple utilisés aujourd'hui pour affiner les prévisions de croissance économique d'une région.

De manière générale, les applications de l'analyse textuelle et de la vision informatique sont très variées et touchent à tous les secteurs : contrôle des processus industriels, médical, navigation, sécurité, environnement, art etc. Une bonne connaissance de ces systèmes est donc indispensable pour mieux appréhender l'évolution de nos économies.

Compétences à acquérir :

Création et numérisation de corpus Calculs de similarité et analyse de sentiment Deep learning Manipulations avec R & Python.

Pré-requis obligatoires

Connaissances de base de R et Python. Appétence pour la programmation.

Mode de contrôle des connaissances :

Mémoire

Coefficient : 0.50