02
MAI.
2023

de 12h à 13h
Rockefeller - Amphi A CIER

SÉMINAIRE

Classification de données médicales textuelles et structurées par apprentissage profond

Alexandre Bailly
-
La pratique de la médecine conduit à récolter de nombreuses données, qu'elles soient structurées ou textuelles, prodiguant de nombreuses informations sur le patient et son état. L'analyse de ces données étant souvent fastidieuse, des modèles d'Intelligence Artificielle tels que des Réseaux de Neurones peuvent être utilisés notamment pour le développement de modèles de classification. Entraînés par des méthodes d'apprentissage profond, les Réseaux de Neurones requièrent plus de données labellisées que des modèles classiques comme les Régressions Logistiques mais sont théoriquement capables d'approximer n'importe quelle fonction. L'apprentissage profond est notamment très utile pour la classification de textes en Traitement Automatique de la Langue Naturelle, en particulier avec les modèles de langues contextualisés. La labellisation des textes étant une tâche chronophage, des méthodes d'apprentissage semi-supervisées permettent d'utiliser également des données non labellisées lors de l'entrainement. Diverses formes de données pouvant être récoltées, leur utilisation simultanée au sein d'un modèle de classification requiert l'utilisation de modèles multimodaux. Cela permet théoriquement d'extraire plus d'information conduisant par conséquent à une meilleure classification.