

Informatique

Data Mining et Machine Learning (36 heures)

26_27_4IRC_08_COMSC_UE9_090_MAJ

ACQUIS

- Identifier les anomalies présentes dans les données, repérer les patterns ou tendances, et comprendre les enjeux liés à la construction de modèles de données.
- Concevoir un système de recommandation adapté aux besoins spécifiques.
- Analyser une problématique, sélectionner une méthode appropriée de traitement des données, et l'appliquer efficacement pour obtenir des résultats pertinents.

CONTENU

- Représentation, manipulation et prétraitement des données
- Comprendre les différentes représentations de données : bases relationnelles, NoSQL, données structurées et non structurées.
- Manipuler les données via des API et effectuer leur prétraitement (nettoyage, gestion des valeurs manquantes, transformation).
- Atelier pratique : Manipulation et nettoyage des données avec Pandas, NumPy et Jupyter.
- Traitement des données
- Détecter les anomalies dans les ensembles de données.
- Extraire des connaissances à l'aide de règles d'association.
- Mettre en oeuvre des techniques de partitionnement (clustering) et des approches de classification et régression pour la modélisation prédictive.
- Atelier pratique : Application des méthodes de traitement avec Scikit-learn et Jupyter.
- Construction de modèles et apprentissage machine
- Introduction aux paradigmes d'apprentissage : supervisé, non supervisé, semi-supervisé et par renforcement.
- Construire et entraîner des modèles prédictifs simples.
- Atelier pratique : Création de modèles avec Scikit-learn et introduction aux réseaux de neurones.
- Réseaux de neurones artificiels et apprentissage profond
- Étudier la structure des réseaux (couches, noeuds, fonctions d'activation) et leurs algorithmes d'apprentissage (rétropropagation, descente de gradient).
- Appliquer des techniques d'optimisation et de régularisation pour améliorer les performances.
- Introduction aux concepts fondamentaux de l'apprentissage profond (Deep Learning).
- Projet final
- Intégrer l'ensemble des notions étudiées dans un projet appliqué, allant de la préparation des données à la construction et l'optimisation de modèles prédictifs.

PRÉREQUIS

- Connaissance en langage de programmation Python
- Connaissance de la ligne de commande

PÉDAGOGIE

ÉVALUATION

BIBLIOGRAPHIE