Contenus

Code du Cours	Nom du Cours	Semestre du Cours	Cours Théoriques	Travaux Dirigés (TD)	Travaux Pratiques (TP)	Crédit du Cours	ECTS
VM 536	Applications de science des données	3	0	4	0	3	8

Cours Pré-Requis	
Conditions d'Admission au Cours	

Langue du Cours	Anglais
Type de Cours	Obligatoire
Niveau du Cours	Master
Objectif du Cours	L'objectif du cours est d'apprendre à gérer, avec une discipline d'ingénierie, l'ensemble du cycle de vie des modèles d'apprentissage automatique, du développement à la mise en production, et d'introduire la construction de systèmes fondés sur les grands modèles de langage.
Contenus	Le cours traite des pratiques d'ingénierie nécessaires pour mettre en production des modèles de science des données et d'apprentissage automatique : gestion de versions, tests du code et des données, ingénierie et sélection des caractéristiques, principes du MLOps et gestion des expériences, automatisation des pipelines par intégration continue, déploiement des modèles par API et conteneurs, surveillance des modèles en production et livraison continue. La dernière partie est consacrée aux grands modèles de langage : fondements, ingénierie des prompts, techniques avancées et architectures de systèmes fondées sur les LLM (LLMOps et agents).
Ressources	Page du cours : https://gitlab.com/onayg/vm536 Foundations of Data Science: Avrim Blum, John Hopcroft, and Ravindran Kannan

Intitulés des Sujets Théoriques

Semaine	Intitulés des Sujets
1	Outils Git et gestion de versions
2	Stratégies de test pour le code et les données
3	Ingénierie et sélection des caractéristiques
4	Principes du MLOps et gestion des expériences
5	Automatisation des pipelines par intégration continue (CI)
6	Stratégies de déploiement des modèles : API et conteneurs
7	Examen de mi-parcours
8	Surveillance des modèles en production et livraison continue (CD)
9	Fondements des grands modèles de langage (LLM) et ingénierie des prompts
10	Techniques avancées pour les LLM
11	Architectures de systèmes fondées sur les LLM : LLMOps et agents