Contenus

Code du Cours	Nom du Cours	Semestre du Cours	Cours Théoriques	Travaux Dirigés (TD)	Travaux Pratiques (TP)	Crédit du Cours	ECTS
INF 511	Science des Données	1	3	0	0	3	6

Cours Pré-Requis	
Conditions d'Admission au Cours	

Langue du Cours	Anglais
Type de Cours	Électif
Niveau du Cours	Master
Objectif du Cours	Ce cours vise à introduire les étudiants au processus d'exploration de données. Parmi les principaux objectifs du cours figurent la compréhension et l'utilisation des techniques de préparation et de prétraitement des données, de divers algorithmes d'exploration de données ainsi que des outils permettant d'évaluer leurs résultats. Le cours se concentre sur les approches standards liées à l'exploration des règles d'association, à la classification supervisée et à la classification non supervisée (clustering). Des connaissances statistiques de base sont nécessaires pour comprendre les algorithmes d'exploration et les outils d'évaluation de la qualité. Ainsi, le cours vise à permettre aux étudiants de produire des solutions pratiques dans le domaine de l'analyse des données.
Contenus	1. Introduction à l'exploration de données et à l'analytique prédictive 2. Prétraitement des données, analyse exploratoire des données 3. Méthodes de réduction de dimension, analyse statistique univariée 4. Statistiques multivariées, préparation des données pour la modélisation 5. Régression linéaire simple, régression multiple 6. Construction de modèles 7. Algorithme des k plus proches voisins, arbres de décision 8. Régression logistique, Naive Bayes et réseaux bayésiens 9. Examen de mi-semestre 10. Techniques d'évaluation des modèles 11. Évaluation graphique des modèles de classification 12. Clustering hiérarchique et k-means, mesure de la qualité des clusters 13. Règles d'association, méthodes d'ensemble (ensemble methods) 14. Présentations des étudiants
Ressources	1. Data Mining - Practical Machine Learning Tools, 2nd edition, Ian H. Witten & Eibe Frank, Morgan Kaufmann, 2005. 2. Neural Networks - A Comprehensive Foundation, 2nd edition, Simon Haykin, Pearson/Prentice Hall,1999. 3. Data Mining: Concepts and Techniques, Jiawei Han & Micheline Kamber, Morgan Kaufmann, 2000. 4. Applied Statistics and Probabilities for Engineers, 4th edition, D.C. Montgomery & G.C. Runger, John Willey & sons, 2006. 5. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edition, T. Hastie, R. Tibshirani & J. Friedman, Springer, 2009.

Intitulés des Sujets Théoriques

Semaine	Intitulés des Sujets
1	Introduction à l'exploration de données et à l'analytique prédictive
2	Prétraitement des données, analyse exploratoire des données
3	Méthodes de réduction de dimension, analyse statistique univariée
4	Statistiques multivariées, préparation des données pour la modélisation
5	Régression linéaire simple, régression multiple
6	Construction de modèles
7	Algorithme des k plus proches voisins, arbres de décision
8	Régression logistique, Naive Bayes et réseaux bayésiens

Semaine	Intitulés des Sujets
9	Examen partiel
10	Techniques d’évaluation des modèles
11	Évaluation graphique des modèles de classification
12	Clustering hiérarchique et k-means, mesure de la qualité des clusters
13	Règles d’association, méthodes d’ensemble (ensemble methods)
14	Présentations des étudiants