Contenus

	Nom du Cours		Semestre du Cours	Cours Théoriques	Travaux Dirigés (TD)	Travaux Pratiques (TP)	Crédit du Cours	ECTS
1	NF 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	Le but de ce cours est d'introduire le processus de fouille de données aux étudiants. Ceci inclut la description des tâches de préparation des données et de pré-traitement, les algorithmes de fouille classiques, et les outils disponibles pour évaluer leurs résultats. Le cours se concentre sur les approches standardpour ce qui concerne la recherche de règles d'association, et la classification supervisée et non-supervisée. Des bases en statistiques sont nécessaires pour comprendre les algorithmes étudiés et les méthodes d'évaluation.
Contenus	 Introduction Préparation des données Règles d'association et algorithme a priori FP-trees et règles complexes Arbres de décision et classificateur de bayes naif Régression statistique et réseaux bayesiens Réseaux de neurones et autres classificateurs Évaluation de la qualité de la classification Comparaison de classificateurs Distance et partitionnement Méthode de partitionnement hiérarchiques Partitionnement par grille et par densité Traitement à base de modèles
Ressources	 Data Mining - Practical Machine Learning Tools, 2nd edition, Ian H. Witten & Eibe Frank, Morgan Kaufmann, 2005. Neural Networks - A Comprehensive Foundation, 2nd edition, Simon Haykin, Pearson/Prentice Hall,1999. Data Mining: Concepts and Techniques, Jiawei Han & Micheline Kamber, Morgan Kaufmann, 2000. Applied Statistics and Probabilities for Engineers, 4th edition, D.C. Montgomery & G.C. Runger, John Willey & sons, 2006. 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	
2	Préparation des données	
3	Règles d'association et algorithme a priori	
4	FP-trees et règles complexes	
5	Arbres de décision et classificateur de bayes naif	

Semaine	Intitulés des Sujets	
6	Régression statistique et réseaux bayesiens	
7	Réseaux de neurones et autres classificateurs	
8	Évaluation de la qualité de la classification	
9	Comparaison de classificateurs	
10	Distance et partitionnement	
11	Méthode de partitionnement hiérarchiques	
12	Partitionnement par grille et par densité	
13	Traitement à base de modèles	
14	Présentation des travaux bibliographiques	