

Data Mining			13X011	
Alexandros KALOUSIS (CC)				
Nombre d'heures par semaine	cours	2	Semestre d'automne	
	exercices	2	Semestre de printemps	<input checked="" type="checkbox"/>
	pratique		Total d'heures	56
Cursus		Type		Crédits ECTS
Bachelor en sciences informatiques		Obligatoire		4

OBJECTIFS :

- Comprendre les tâches de base de l'exploration de données
- Comprendre les principes de base de l'apprentissage supervisé
- Comprendre les principaux concepts des algorithmes d'apprentissage : fonctions objectifs, familles de modèles et structure de modèles, complexité des modèles, optimisation et régularisation des modèles
- Être capable d'opérationnaliser (c'est-à-dire implémenter et d'utiliser) les principaux concepts d'apprentissage automatique d'une manière qui corresponde aux exigences des problèmes du monde réel.
- Acquérir une bonne méthodologie d'expérimentation et d'évaluation

CONTENU :

Introduction à l'exploration de données et à l'apprentissage automatique
 Erreur de Bayes et classificateur de Bayes
 Approches d'inférence : maximisation de vraisemblance et maximum a posteriori
 L'approche bayésienne
 Comparaison et sélection de modèles
 Estimation de la densité

- Méthodes du voisin le plus proche, métriques, distance de Mahalanobis et ses propriétés
- Estimation de la densité basée sur les noyaux

Algorithmes discriminants linéaires

- Transformations linéaires, métriques et interprétation géométrique
- Approches génératives
- Approches discriminatoires
 - Régression logistique
- Approches de fonction discriminante
 - Perceptron
- Discriminants linéaires de Fisher

Régression linéaire

- Régularisation, l_1 et l_2
- Interprétation géométrique
- subgradients

Introduction aux réseaux de neurones

- L'algorithme de rétropropagation
- Graphes computationnelles

Méthodologie d'évaluation

- Méthodes de rééchantillonnage
- Test de significativité

Forme de l'enseignement

2h cours et 2h exercices / implémentation des algorithmes / préparation du projet DM.

Documentation	Ouvrages de références
Préalable requis	Statistiques et Probabilités, Algèbre, Calculs
Préparation pour	-
Mode d'évaluation	Exercices à proposer et à remettre en séance de TP. Un projet à remettre avant la fin du semestre. La note finale est la moyenne pondérée des notes des exercices (40%) et du projet (60%). Elle est arrondie au demi-point le plus proche.
Sessions d'examens	J/AS