

CERTIFICAT

SCIENCES DES DONNÉES ET BIG DATA OUTILS ET INTRODUCTION



DATE	PROGRAMME*	FORMAT	HORAIRES	LIEU
Jeudi 10 septembre 2020	Réunion de rentrée	En ligne		en ligne
MODULE DE SENSIBILISATION				
Du 11 au 30 septembre 2020	M1.1 Introduction à R / <i>Exercice de manipulation de variables aléatoires et leurs distributions statistiques</i> M1.2 Elements d'analyse statistique de données : statistique élémentaire, descriptive unidimensionnelle et bidimensionnelle, estimation et tests statistiques/ <i>Autoévaluation avec jeu de données du pic d'ozone de la ville de Rennes</i> M1.3 Modélisation statistique : modèle gaussien-régression linéaire multiple, modèle binomial-régression logistique) / <i>Autoévaluation avec prédiction du pic d'ozone de la ville de Rennes</i>	E-learning	-	-
Jeudi 1er octobre 2020	Séance de discussion questions/réponses sur les contenus de la première phase d'e-learning	Présentiel	14h30-16h30	INSA, Vinci
Du 2 octobre au 25 octobre 2020	M1.4 Introduction à Python M1.5 Exploration multidimensionnelle : data mining, ACP-AFD, clustering avec k-means / <i>Tutoriel en autonomie : intro ACP-AFD, adaptation statistique et prévision de la concentration en ozone (exploration) ; segmentation d'image et cartographie de la surface de Mars</i> M1.6 Principes de l'apprentissage statistique / <i>Tutoriel en autonomie : classification supervisée de données synthétiques dans R² et adaptation statistique et prévision de la concentration en ozone (apprentissage et prévision)</i>	E-learning	-	-
Lundi 26 octobre 2020	TP sur l'exploration multidimensionnelle / tutoriel : reconnaissance d'activité à partir des enregistrements d'un smartphone (exploration)	Présentiel	9h30-12h30	INSA, 101 INSA, 102 INSA, 103 en ligne
	TP sur l'apprentissage statistique / tutoriel : reconnaissance d'activité à partir des enregistrements d'un smartphone (apprentissage et reconnaissance)	Présentiel	14h-17h	INSA, 101 INSA, 102 INSA, 103 en ligne
Mardi 27 octobre 2020	M1.7 Séance sur les algorithmes classique d'optimisation (introduction sur l'optimisation, algorithmes classiques avec gradient et spark)	En ligne	9h-11h	en ligne
	TP programmer 1 algorithme	En ligne	11h-13h	en ligne
Mercredi 28 octobre 2020	M1.8 Séance sur les techniques de virtualisation et containerisation; les plateformes cloud	En ligne	9h-11h	en ligne
	TP Déploiement dans une infrastructure virtualisée	En ligne	11h-13h	en ligne
Du 6 novembre 2020 au 31 janvier 2021	M1.9 Projet de fin de module par binômes "Développer un algorithme sur une infrastructure"	Projet		
MODULE IMMERSION				
Mardi 3 novembre 2020	M2.1 Séance sur les infrastructures logicielles pour le big data: MapReduce, démystification Hadoop et Spark	En ligne	18h-20h	-
Jeudi 5 novembre 2020	TP avec Spark	En ligne	14h-18h	en ligne
Jeudi 19 novembre 2020	M2.2 Séance d'introduction aux méthodes d'apprentissage : arbre binaire de décision, réseaux de neurones, support vector machines, agrégation d'arbres	Présentiel	14h-18h	ISAE-SUPAERO
Jeudi 3 décembre 2020	TP Arbre binaire de décision	Présentiel	14h-17h	ISAE-SUPAERO
Du 4 décembre 2020 au 6 janvier 2021	M2.3 Les méthodes d'apprentissage : arbre binaire de décision, réseaux de neurones, support vector machines, agrégation d'arbres	E-learning	-	-
Jeudi 7 janvier 2021	TP spam	Présentiel	14h-17h	ISAE-SUPAERO
Jeudi 11 février 2021	M2.4 Séance sur les algorithmes stochastiques plus sophistiqués: optimisation parcimonieuse, factorisation non négative de matrice	Présentiel	14h-16h	INP-ENSEEIH
	TP programmation d'un algorithme de chaque type	Présentiel	16h-18h	INP-ENSEEIH
Du 12 février 2021 au 31 mars 2021	M2.5 Projet de fin de module par équipes de 4 : challenge Kaggle	Projet	-	en ligne
CONFERENCES				
Mardi 13 octobre	Cas d'application dans l'industrie	Présentiel	18h30-20h30	INP-ENSEEIH, B00
Mardi 10 novembre	Cas d'application dans l'industrie	Présentiel	18h30-20h30	INP-ENSEEIH, B00
Mardi 15 décembre	Ouverture	Présentiel	18h30-20h30	INP-ENSEEIH, B00
Mardi 19 janvier	Cas d'application dans l'industrie	Présentiel	18h30-20h30	INP-ENSEEIH, B00
Mardi 16 février	Ouverture	Présentiel	18h30-20h30	INP-ENSEEIH, B00

*ce programme est susceptible d'être modifié