

SYLLABUS DU COURS

INFORMATIONS GÉNÉRALES

Faculté : Faculté des sciences Département : Département d'Informatique Intitulé : Machine Learning Code : Semestre : Semestre 6	Nombre de Crédits : Langue d'enseignement : Anglais Prérequis : Horaire : Salle :
Enseignant(e) : Mira ABBOUD Téléphone : 71 68 96 36 Courriel : mira.abboud90@gmail.com	Adresse bureau : Heures de visites :

INFORMATIONS GÉNÉRALES

Lien avec le référentiel de compétences / les résultats d'apprentissage programme :

Ce cours familiarisera les étudiants avec un large éventail de modèles et d'algorithmes pour l'apprentissage automatique et préparera les étudiants à la recherche et/ou à l'application industrielle des techniques d'apprentissage automatique.

Objectif général ou finalité :

Ce cours présente tous les modules importants qu'un scientifique de données doit connaître, y compris l'apprentissage automatique et le langage de programmation python. Il enseigne également des concepts importants tels que l'acquisition de données, l'exploration de données, le traitement de données et l'analyse de données. Ce cours est conçu en gardant à l'esprit les tendances actuelles de l'industrie et les compétences requises pour devenir un scientifique des données à succès.

RÉSULTATS D'APPRENTISSAGE- RAUE

❖ Résultats des connaissances fondamentales

Ce cours vise à fournir aux étudiants / permet aux étudiants de:

- 1) Ecrire des programmes en Python
- 2) Comprendre les problèmes résolubles avec la data science
- 3) Travailler selon le processus de KDD
- 4) Collecter, manipuler, mélanger les données de différentes sources de données
- 5) Visualiser les données et effectuer une analyse exploratoire des données
- 6) Comprendre le cycle de vie d'un projet de data science
- 7) Créer des produits pour les applications d'entreprise
- 8) Obtenir les bases des techniques d'apprentissage automatique supervisé et non supervisé.

❖ Résultats d'application

Ce cours forme la base de presque tous les autres cours liés à la data science et l'intelligence artificielle. Une bonne compréhension de ce cours et des algorithmes est essentielle pour tous les autres sujets avancés.

❖ Résultats d'intégration

Apprenez comment et quand utiliser différents algorithmes de data science.

❖ Résultats d'apprentissage-comment-apprendre

L'objectif principal sera d'inciter les élèves à aimer la matière. Une fois que cela se produira, les étudiants s'efforceront seuls pour le reste de leur carrière d'apprendre par eux-mêmes.

Contenu, Méthodes et Séances

Résultats d'apprentissage	Séances	Contenu	Méthodes d'enseignement
Introduction to Machine Learning	1	Introduction to the Knowledge Discovery from Databases + KDD process + example of applications	Motivational course (ppt)
Python basics	1	Jupyter notebook + python basics (statements, import packages, dataframe, set, dictionaries)	Course (ppt) + small applications on python
Data preparation + exploration	1	Read/Write on CSV, excel, data visualization, data correlation, outliers, describe data	Course (ppt) + Practice Python (read, dataframes, describe, plotting methods)
Data pre-processing	2	Inconsistent values, duplicate records, missing values, invalid data, outliers detection, features selection, scaling, transformation	Course (ppt) + Practice Python (filters, nan values, standardization, fit transform methods)
Regression	2	Intro to supervised learning and regression + regression algorithms (Simple linear regression, multiple linear regression, polynomial regression, SVR (SVR intuition) + SVR non-linear regression)	Course (ppt) + Practice Python (Sklearn library, train and fit the various models on salary database, predict and check the accuracy, SVR module, RBF kernel)
Classification	2	Intro to classification + classification algorithms (Logistic Regression, Decision trees, Random Forest, Nearest Neighbor, SVM)	Course (ppt) + Practice Python (train and fit the various models on social network ads database, LogisticRegression, DecisionTreeClassifier, RandomForestClassifier ,

			SVC and BF Kernel, KNeighborsClassifier)
Unsupervised learning (clustering + dimensionality reduction)	2	Intro to unsupervised learning + clustering algorithms (K-means, Hierarchical clustering) + Dimensionality Reduction (PCA)	Course (ppt) + Practice Python (KMeans, AgglomerativeClusterin, PCA on Mall customer and Iris databases)
Association Rules	1	Intro to Association Rules + Association algorithms (apriori + Eclat)	Course (ppt) + Practice Python (apriori on Market Basket Optimization database)
Text mining	2	Text preprocessing (tokenizing , Stemming, chunking, chunking, lemmatizing), text classification (application on sentiment analysis) , text clustering (application on clustering tweets)	Course (ppt) + practice Python (NLTK package, corpora, sentiment analysis on Movies review database and clustering on twitter database)

MODALITÉS D'ÉVALUATION

Class attendance	5%
Class participation	10%
Midterm exam	30%
Final exam	55%

RÉFÉRENCES BIBLIOGRAPHIQUES

Géron, Aurélien. <i>Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems</i> . O'Reilly Media, 2019.
Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. <i>The elements of statistical learning</i> . Vol. 1. No. 10. New York: Springer series in statistics, 2001.
Kirk, Matthew. <i>Thoughtful Machine Learning with Python: A Test-Driven Approach</i> . "O'Reilly Media, Inc.", 2017.