

	<b>Fiche module</b>		Mise à jour :
	Cycle de formation d'ingénieurs en Informatique		Page 1 sur 4

<b>Module : Machine Learning</b>			Code
			ING-4-GLSI-S8-P2
Période	<b>S8-P2</b>	Volume horaire	<b>21 h</b>
		ECTS	<b>2</b>

<i>Responsable</i>	<b>Mr. Nadjib Bendaoud</b>	<i>email</i>	nadjib.bendaoud@tek-up.tn
<i>Equipe pédagogique</i>	Mr. Nadjib Bendaoud		

### 1. Objectifs de Module (*Savoirs, aptitudes et compétences*)

Ce module porte sur les algorithmes de Machine Learning.

#### Acquis d'apprentissage :

A la fin de cet enseignement, l'élève sera capable de :

<b>Compétences</b>
<b>C1.2</b> Maitriser les différents algorithmes de Machine Learning.
<b>C1.3</b> Identifier les limites et les avantages de chaque algorithme
<b>C1.4</b> Caractériser et analyser et justifier choix du modèle à développer.
<b>C3.2</b> Avoir les compétences pour résoudre un problème de l'intelligence artificielle.

### 2. Pré-requis (*autres UE et compétences indispensables pour suivre l'UE concernée*)

- Concepts mathématiques
- Des connaissances Physique

### 3. Répartition d'Horaire de Module

<i>Intitulé de l'élément d'enseignement</i>	<i>Total</i>	<i>Cours</i>	<i>TD</i>	<i>Atelier</i>	<i>PR</i>
Module : Machine Learning	21	15		6	

### 4. Méthodes pédagogiques et moyens spécifiques au Module

(*pédagogie d'enseignement, ouvrages de références, outils matériels et logiciels*)

- Supports de Cours
- Projecteur et Tableau
- Travaux dirigés

### Bibliographie

<i>Titre</i>	<i>Auteur(s)</i>	<i>Edition</i>
Pattern Recognition and Machine Learning	Christopher M. Bishop	

The Elements of Statistical Learning: Data Mining, Inference, and Prediction	Trevor Hastie, Robert Tibshirani, and Jerome Friedman	
Machine Learning	Andrew Ng	Coursera

5. Contenu (Descriptifs et plans des cours / Déroulement / Détail de l'évaluation de l'activité pratique)	Durée allouée
<b>Module 1: Machine Learning</b>	
<b>Séance 1 : Introduction à l'IA et au Machine Learning</b>	Cours 1.5 H
<ul style="list-style-type: none"> <li>• Relation entre IA, ML et DL</li> <li>• Apprentissage supervisé</li> <li>• Apprentissage non supervisé</li> <li>• Apprentissage semi supervisé</li> <li>• Apprentissage par renforcement</li> </ul>	
<b>Séance 2 : Régression Linéaire</b>	Cours 1.5 H
<b>Objectifs :</b> <ul style="list-style-type: none"> <li>• Comprendre le modèle de régression linéaire.</li> <li>• Apprendre les concepts théoriques et mathématiques derrière la régression linéaire.</li> </ul> <b>Contenu :</b> <ul style="list-style-type: none"> <li>• Théorie de la régression linéaire.</li> <li>• Hypothèses du modèle de régression linéaire.</li> <li>• Algorithme de descente de gradient pour la régression linéaire.</li> <li>• Régularisation (Ridge, Lasso).</li> <li>• Cas d'utilisation de la régression linéaire.</li> </ul>	
<b>Séance 3 : K-Nearest Neighbors (KNN)</b>	Cours 1.5 H
<b>Objectifs :</b> <ul style="list-style-type: none"> <li>• Comprendre les concepts de base de l'algorithme KNN.</li> <li>• Apprendre les principes théoriques derrière KNN.</li> </ul> <b>Contenu :</b> <ul style="list-style-type: none"> <li>• Introduction à l'algorithme KNN.</li> <li>• Distance métrique (Euclidienne, Manhattan, Minkowski).</li> <li>• Choix du paramètre K.</li> <li>• Cas d'utilisation de KNN.</li> </ul>	

**Séance 4 : Arbres de Décision**
**Objectifs :**

- Comprendre les concepts de base des arbres de décision.
- Apprendre les différentes implémentations des arbres de décision

**Contenu :**

- Introduction aux arbres de décision.
- Différents algorithmes d'arbres de décision (CART, ID3 ..)
- Critères de division des nœuds d'un arbre (Gini, Entropy).
- Techniques d'élagage d'arbre
- Cas d'utilisation des arbres de décision.

Cours

3H

**Séance 5 : Data Engineering, validation et sélection des modèles**

Cours

3H

Modèles paramétrique vs non paramétriques

- Paramètres vs Hyper-paramètres
- Optimisation des paramètres
- Optimisation des hyper-paramètres

Data Engineering et préparation des données :

- Pré-traitement des données (nettoyage, transformation, normalisation).
- Ingénierie des features (sélection et extraction).
- Encodage des variables catégorielles.

**Séance 6: TP**

Cas d'utilisation pour la regression

- Préparation du jeu de données
- Régression Linéaire
- K-NN
- Arbre de décision

Atelier

3H

Cas d'utilisation pour la classification

- Préparation du jeu de données
- Régression Linéaire
- K-NN
- Arbre de décision

**Séance 10 : Ensemble Methods**

Cours

1.5H

- Apprendre la théorie derrière les techniques de Bagging et Boosting.
- Introduction aux ensembles de modèles.
- Bagging (Bootstrap Aggregating) et Random Forests.
- Boosting, AdaBoost et Gradient Boosting.

**Séance 13 : Apprentissage non-supervisé : Clustering**
**Objectifs :**

- Comprendre le principe du clustering.
- Connaitre des algorithmes comme DBSCAN et K-means.

**Contenu :**

- Introduction au clustering.
- Algorithmes de clustering (K-Means, DBSCAN, Hierarchical Clustering, elbow method).
- Évaluation du clustering (silhouette score).
- Cas d'utilisation du clustering dans la vie réelle.

Cours 3H

**Séance 14 :**

- Application des Random Forests
- Utilisation du Gradient Boosting
- Découverte des modèles XGBoost et CatBoost pour une meilleure précision
- Cas d'utilisation Apprentissage non supervisé
- Clustering en utilisant l'algorithme du K-means

Atelier 3 H

**6. Mode d'évaluation de Module**(nombre, types et pondération des contrôles)

Eléments d'enseignement	Coeff	DS	EX	TP	PR
Module – Machine Learning	1	0%	60%	0%	40%

Pour valider le module, les étudiants passeront un examen dont le coefficient est de 60%, et un PR dont le coefficient est de 40%.

La durée de tous les examens (Examen, DS...) est de 1h30.

Quant à l'examen, il est planifié après l'écoulement des 7 semaines et portera sur toutes les thématiques enseignées tout au long des 21 heures.

Concernant le PR, il est planifié une semaine avant l'examen.

Le module est validé si l'étudiant obtient une moyenne supérieure ou égal à 10 sur 20.