



Université Mohamed Premier Oujda  
École Nationale de l'Intelligence Artificielle et du Digital Berkane  
Année universitaire : 2024 / 2025

Filière : 1A  
Prof : Mohamed Khalifa BOUTAHIR

## MACHINE LEARNING II – TP 1

### Objectif :

L'objectif de ce TP est de se familiariser avec les outils essentiels du Reinforcement Learning (RL), notamment OpenAI Gym. Les étudiants vont explorer comment interagir avec un environnement RL et exécuter des actions avant d'implémenter un algorithme d'apprentissage dans les séances suivantes.

### A) Partie 1 : Présentation des Bibliothèques Clés

#### 1. OpenAI Gym

- OpenAI Gym est une bibliothèque permettant de simuler des environnements interactifs pour tester des algorithmes de RL.
- Un environnement Gym est défini par un ensemble d'états, d'actions, de récompenses et d'un critère de fin.

#### ★ Installation de Gym

```
pip install --upgrade gymnasium pygame numpy
```

#### ★ Création d'un environnement

```
import gymnasium as gym

env = gym.make("CartPole-v1", render_mode="human")
env.reset()
```

### B) Partie 2 : Exercices Pratiques avec OpenAI Gym

- Exercice 1 : Découverte et Exploration d'un Environnement Gym

**Objectif :** Comprendre la structure d'un environnement Gym en explorant ses propriétés et ses actions possibles.

#### ✓ Instructions :

1. Afficher l'espace d'actions et l'espace d'observations.
2. Exécuter une boucle de simulation avec des actions aléatoires pendant 100 itérations.



### 3. Observer les valeurs des observations retournées.

```
print(f"Espace d'actions : {env.action_space}")
print(f"Espace d'observations : {env.observation_space}")

for _ in range(100):
    action = env.action_space.sample()
    observation, reward, done, _, _ = env.step(action)
    print(f"Action : {action}, Observation : {observation}, Reward : {reward}")
    if done:
        env.reset()

env.close()
```

#### • Exercice 2 : Manipulation des Observations et Récompenses

**Objectif :** Comprendre comment récupérer les observations et les récompenses lors de l'interaction avec l'environnement.

✓ Instructions :

1. Prendre une action et récupérer les valeurs retournées (observation, reward, done).
2. Afficher ces valeurs et analyser leur signification.
3. Faire plusieurs essais et noter les variations des observations et des récompenses.

#### • Exercice 3 : Contrôle Manuel de l'Agent

★ **Objectif :** Permettre à l'utilisateur de contrôler manuellement l'agent pour mieux comprendre l'effet des actions.

✓ Instructions :

1. Demander à l'utilisateur d'entrer une action (0 ou 1).
2. Exécuter l'action dans l'environnement et afficher les nouvelles observations.
3. Répéter l'opération jusqu'à la fin de l'épisode.
4. Afficher la durée totale de l'épisode avant qu'il ne se termine.

#### • Exercice 4 : Évaluation des Performances d'une Politique Aléatoire

★ **Objectif :** Mesurer la durée moyenne d'un épisode lorsqu'un agent prend des actions aléatoires.

✓ Instructions :

1. Faire exécuter à l'agent des actions aléatoires pendant plusieurs épisodes (ex : 10 épisodes).
2. Calculer la durée moyenne avant que l'épisode ne se termine.
3. Comparer les résultats entre plusieurs exécutions.