

# Bandits

TP À RENDRE LE 24/09/25 23h59

## 1 Présentation du problème

Nous allons appliquer les algorithmes de bandits vus lors de la séance précédente. Pour cela on utilisera le code `forTP1.py` qui propose une classe simplifiée de Bandits avec laquelle on peut interagir par le biais de la fonction `get_arm`.

## 2 Objectifs

Votre objectif est à l'aide de chacun des algorithmes vus en cours (en dehors celui par "ascension" de gradient) de proposer une estimation du meilleur bras pour diverses expériences (variation de  $K$ , de  $T$ ).

Pour rappel, ces algorithmes sont :

- Uniform Exploration
- $\epsilon$ -greedy
- Successive elimination
- UCB1

Pour cela, vous aurez peut-être à fixer certains paramètres des méthodes, vous indiquerez dans le compte-rendu les valeurs prises en les justifiant si possible.

Ensuite, vous devrez proposer sur l'un des algorithmes une méthodologie empirique d'estimation du regret, que vous comparerez aux formules données dans le cours.

## 3 Rendus

- Un Notebook.