

# APPRENTISSAGE PAR RENFORCEMENT

## M2 DAC

### TME 1. Problèmes de Bandits

Ce TME a pour objectif d'expérimenter les modèles UCB et LinUCB vus en cours pour de la sélection de publicité en ligne.

#### 1 Données

Un fichier de données pour la sélection de publicité en ligne a été préparé pour expérimenter les modèles: CTR.txt, téléchargeable sur le site de l'UE. Ce fichier contient les contextes (profils d'articles) et les taux de clic sur les publicités de 10 annonceurs pour 5000 articles (1 par ligne):

*<numero de l'article>:<représentation de l'article en 5 dimensions séparées par des ";">:<taux de clics sur les publicités de 10 annonceurs séparés par des ";">*

Pour chaque visite, l'objectif est de choisir la publicité d'un des 10 annonceurs permettant d'engranger le plus fort taux de clics.

#### 2 Baselines

Trois baselines peuvent être envisagées:

- Stratégie Random : A chaque itération, on choisit n'importe quel annonceur
- Stratégie StaticBest : A chaque itération, on choisit l'annonceur avec le meilleur taux de clics cumulés
- Stratégie Optimale : A chaque itération, on choisit l'annonceur qui a le meilleur taux de clics à cette itération

#### 3 UCB

Développer et expérimenter l'algorithme UCB vu en cours

#### **4 LinUCB**

Développer et expérimenter l'algorithme LinUCB vu en cours. Cet algorithme propose de prendre en compte le contexte de la décision, qui peut être très utile pour prévoir les variations des taux observés.