

PROJET DE REINFORCEMENT LEARNING

Réalisé par :

Hinault EZOBA, Harouna OUILY, Kodjo EDEM, Koami AZIABOU

Sujet : « Fixation de prix dans un marché concurrentiel avec reinforcement learning »

1. Présentation du Processus de décision de Markov (MDP) : approche théorique

1. États (S):

L'ensemble des états possibles du système représente toutes les combinaisons des variables pertinentes pour la prise de décision de prix. Dans ce cas, les états pourraient être définis par :

- `prix_entreprise` : Le prix de vente fixé par l'entreprise à l'état actuel
- `prix_concurrents` : Le prix de vente moyen des concurrents dans le marché à l'état actuel
- `demande_marche` : Le niveau de la demande dans le marché à l'état actuel
- `stock_marche_moyen` : Le niveau de stock dans le marché à l'état actuel
- `cout_entreprise` : Le coût de revient de l'entreprise pour le produit à l'état actuel
- `stock_entreprise` : Quantité de biens en stock dans l'entreprise à l'état actuel
- `Qt_vendue_entreprise` : Quantité de stock vendue par l'entreprise à l'état actuel
- `date` : Date de l'observation à l'état actuel

Par exemple d'une ligne dans le dataset (voir dataset en attache) dont, un état pourrait être représenté par :

```
(prix_entreprise = 4,87 ; prix_concurrents = 0,59 ; demande_marche = 8952 ;  
stock_marche_moyen = 24425 ; cout_entreprise = 5,19 ; date = 31-12-2023)
```

2. Actions (A) :

L'ensemble des actions possibles représente les choix de prix que l'entreprise peut effectuer.

- Une augmentation du prix
- Une diminution de prix
- Un maintien du prix

3. Fonction de récompense (R)

La fonction de récompense $R(s, a)$ définit la récompense immédiate obtenue par l'entreprise en effectuant une action a dans un état s .

Dans ce problème, la récompense pourrait être définie comme une fonction du profit de l'entreprise calculé comme suit:

$$\text{profit}(s, a) = (PE_s - C_s) * Qt_vendue_entreprise$$

Or, à long terme le prix du marché deviendra constante au fur et à mesure que les nouveaux intervenants entrent dans le marché et la seule variable et une de variable pouvant rassurer le profit de l'entreprise sera la part du marché.

Connaissant la formule de la part du marché :

$$\text{part_marche}(s, a) = Q_{t_vendue_entreprise} / \text{demande_marche}$$

Par conséquent

$$Q_{t_vendue_entreprise} = \text{part_marche}(s, a) * \text{demande_marche}$$

En substituant la quantité vendue dans l'expression du profit on trouve :

$$\text{profit}(s, a) = (PE_s - C_s) * \text{part_marche}(s, a) * \text{demande_marche}$$

$$R(s, a) = \text{profit}(s, a) = (PE_s - C_s) * \text{part_marche}(s, a) * \text{demande_marche}$$

où:

- $\text{profit}(s, a)$: Le profit généré par l'entreprise en fixant le prix a dans l'état s
- $\text{part_marche}(s, a)$: La part de marché obtenue par l'entreprise en fixant le prix a dans l'état s
- C_s : cout_entreprise à l'état S
- PE_s : prix_entreprise à l'état S

2. Présentation du script de Processus de décision de Markov (MDP) et algorithme de renforcement learning

Voir fichier « `Projet_Reinforcement_learning.py` » dans le repository Github

Ce script va produire au final un fichier nommé « `q_table.npy` » qui sera utilisé pour produire une interface de l'utilisateur avec `gradio.py`

3. Présentation du script de l'interface de l'utilisateur

Voir fichier « `Interface_action_prix.py` » dans le repository Github

Ce script va produire une interface pour l'utilisateur auquel en introduisant l'Etat actuel l'application renvoie l'action à réaliser jusqu'à ce que le prix optimal sera déterminé.