

Dernièrement, la présence de commerce sur le monde digital a explosé, la crise sanitaire semble même avoir accéléré cette tendance. C'est dans cette situation d'ultra compétitivité qu'est née notre problématique :

*Quelle stratégie de dynamic pricing à adopter dans un marché en présence de clients stratégiques avec un produit ayant une maturité fixée (finie ou infinie) ?*

Pour apporter une solution à cette problématique nous avons choisi de créer un modèle avec d'un côté un vendeur qui fixe les prix de manière intelligente et de l'autre des clients avec des comportement variés.

Notre solution se propose dans un premier temps de définir deux types de clients : naïf et stratégique. Cette prise en compte du comportement client est un élément déterminant, car ce point n'est pas utilisé par les solutions existantes. Un client naïf a un comportement prévisible et pas de connaissance du marché.

Sa volonté d'achat est parfois régie par le hasard mais nous avons également défini une proportion de clients naïfs qui doivent effectuer un achat, ce qui détermine donc une probabilité pour chaque client d'acheter.

Un client stratégique a, quant à lui, une réflexion sur le marché et l'évolution des prix.

Tout d'abord il a accès à l'évolution historique du prix du produit, ce qui lui permet de baser ses réflexions sur une donnée.

Ensuite, il convient de définir la fourchette de prix acceptables par le client stratégique. Cette fourchette est déterminée par la moyenne des 20% des prix les plus élevés et plus bas. Puis, nous définissons le Willingness-To-Pay qui représente la volonté d'achat du client lorsque le prix est contenu dans une fourchette de prix acceptables. Concrètement, il a fallu définir un signal fort d'achat, en fonction de l'évolution des prix. Ce signal est une baisse par rapport aux deux derniers prix. Si le prix du mois  $M$  est inférieur aux prix des mois  $M-1$  et  $M-2$ , alors l'achat se fait.

Nous modélisons ensuite le vendeur par un réseau de neurones profond approximant la Q-fonction (DQN), ayant l'avantage par rapport au simple Q-learning, de lui transmettre une grande quantité de données.

Les données envoyées au réseau sont l'ensemble des états possibles de l'environnement. Nous avons considéré les états suivants : date, prix et le nombre de réservations, mais nous pourrions en considérer davantage pour aider le réseau à apprendre plus efficacement.

Sur cette base de données, le vendeur apprend les conséquences de ses actions, c'est-à-dire fixer un prix pour un état donné, grâce à une récompense reçue, positive ou négative. Le but du vendeur durant l'apprentissage est alors de maximiser cette récompense à long terme.

Lors de l'apprentissage, le vendeur explore l'environnement d'abord puis exploite ces informations. Les actions prises et la conséquence induite sont enregistrées dans la mémoire du vendeur sous la forme d'une expérience, ainsi toutes les expériences passées peuvent être exploitées afin de maximiser sa récompense.

Une fois le vendeur modélisé, nous le mettons face au marché, donc face aux différentes proportions de clients stratégiques et naïfs. Nous testons sur toutes les proportions possibles de ces deux types de clients et constatons que notre vendeur simulé par le DQN réussit à surperformer sur toutes celles-ci. Pour atteindre ce résultat nous avons entraîné le DQN sur chaque portion de 20% de clients stratégiques. Le DQN est le plus efficace sur la

portion où il y a 80-100% de clients stratégiques ce qui est un résultat très encourageant pour son utilité réelle.

Il nous semble que l'analyse des comportements clients est un aspect important à considérer pour les problématiques de pricing. En effet, cela apporterait une meilleure précision permettant aux entreprises d'obtenir de meilleurs résultats. Concernant les perspectives du projet nous pensons qu'il serait intéressant d'inclure plus d'éléments et également d'optimiser les paramètres d'apprentissage, pour atteindre une précision accrue.