*An adaptive portfolio trading system: A risk-return portfolio optimization using recurrent reinforcement learning with expected maximum drawdown*

L’article traite du développement d’un algorithme d’allocation d’actifs adaptatif, cherchant à optimiser le couple rendement/risque, et utilisant les méthodes d’apprentissage par renforcements récursifs.

Il existe 3 types de méthodes d’apprentissage :

* L’apprentissage non supervisée comme les méthodologies de classification (exemple : regroupement en ACP),
* L’apprentissage supervisé où les individus sont classés en différentes catégories par exemple et l’algorithme essaie de déterminer les règles entre les variables d’entrées permettant de catégoriser au mieux les individus selon les différentes catégories (ex : régression, réseaux neuronaux, etc)
* L’apprentissage par renforcement consistant à faire apprendre à un agent, les décisions à prendre selon un état et afin de maximiser une récompense (ex : robots, voiture autonome, etc)

Les algorithmes par renforcements récursifs en sont une classe particulière, dans la mesure où, comme son nom l’indique, l’algorithme prend en compte ses choix passés, pour définir ses choix futurs.

Nous allons donc essayer de reproduire un système d’allocation dynamique, cherchant à maximiser le ratio de Sharpe entre autres et dont on comparera la performance à celle d’autres méthodologies (Equally-Weighted, etc)

Quelques conseils et astuces :

* Débuter avec un seul actif en utilisant la fonction d’activation *tanh* comme dans la formule (1) dans l’article ;
* Généraliser ensuite dans le cas de plusieurs actifs (et en changeant la fonction d’activation) ;
* Dans les formules (4) et (5), remarquer que :