

Le *Reinforcement Learning* appliqué à la gestion de portefeuille

Dorian Lagadec

12 septembre 2019



université
PARIS-SACLAY

CVISICA

- 1 Introduction
- 2 Aequam Capital : mise en contexte
- 3 Problématique et démarche proposée
- 4 Réalisation et progression du projet
- 5 Pivot, état actuel et étapes à venir

Une année en alternance au sein d'Avisia : plusieurs missions :

- Application de la théorie des graphes à la détection de faux avis en ligne
- Mise en oeuvre d'un scoring de crédit via la plateforme Dataiku
- ...

Année conclue par une **mission longue chez Aequam Capital depuis mi-juin : à la croisée des chemins entre finance de marché et data science.**

- Jeune société de gestion de portefeuille quantitative ;
- Stratégie *long-only* multi-factorielle ;
- Processus d'investissement semi-automatisé en perpétuelle évolution : changement de visages, changement de méthodes, changement d'image.

Dans le cadre de la recherche d'Aequam Capital, la voie qui a été choisie pour orienter la recherche est celle du Reinforcement Learning : **un projet d'exploration pour toutes les parties prenantes.**

- Une logique intuitive et séduisante ;
- Une dimension dynamique et exploratoire appropriée ;
- Un champ d'investigation vaste et innovant (société en phase de développement) ;
- Une littérature prolifique.

Notre problème, en terme de RL : un agent (trader) peut-il, en observant des données qui contiennent supposément un signal (environnement), choisir une allocation de portefeuille parmi plusieurs facteurs (action), afin de maximiser le rendement de long-terme sous des contraintes de volatilité et de perte maximale (récompense) ?

Démarche proposée

- Problématique simple qui cache de nombreuses contraintes opérationnelles et commerciales ;
- Nécessité d'une démarche incrémentale pour permettre une appropriation de la technologie utilisée et du problème concerné.

Démarche envisagée en trois temps (réalisée différemment en pratique) :

- ❶ Exploration d'une stratégie *long-flat* sur un portefeuille équi-pondéré de facteurs ;
- ❷ Exploration d'une stratégie univariée d'un facteur par rapport au portefeuille équi-pondéré ;
- ❸ Exploration d'une stratégie flexible où le poids de chaque facteur dans le portefeuille varie librement entre 0 et 1.

Découverte de l'écosystème du *Reinforcement Learning*

- Revue de littérature : de nombreux articles, **notamment d'application à la finance** (potentiel énorme associé à une complexité avérée) ;
- De nombreux didacticiels qui aident à approcher le fonctionnement des algorithmes sans toutefois fonctionner correctement ;
- Des bibliothèques complètes (Facebook, Google, ...) mais difficiles d'accès pour le néophyte.

- Un choix difficile car pléthorique : de nombreuses données corrélées, peu de données présentant un signal ;
- Les candidats de base : les prix eux-mêmes, les données macroéconomiques (volumes, taux, surprise économique, volatilité, croissance, risque, ...) ;
- Un chantier en soi : la transformation des données - stationnarisation, indicateurs techniques, ...

Implémentation : *trial-and-error*

- Adaptation d'un algorithme et de codes trouvés sur Github ;
- Changements successifs de la fonction - primordiale - de récompense ;
- Ajouts de paramètres spécifiques à notre problème : temps de gel de portefeuille, politique de décision (probabiliste puis déterministe), question de seuil d'indécision ;
- Gestion approximative du problème de *sparse reward* ;
- Tests de couples de paramètres à l'aveugle et visualisation des résultats (performance et apprentissage).

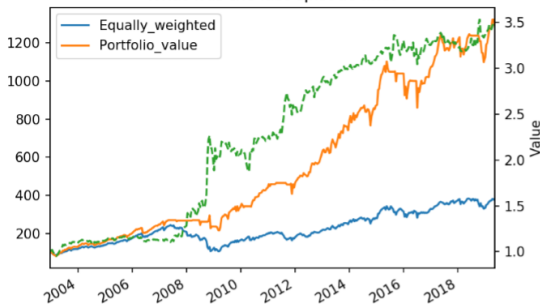
Tout ça a amélioré successivement les performances de l'algorithme, sans toutefois donner de résultats intéressants : une relative impasse...

Etant donné le caractère expérimental du projet et son caractère notoirement difficile, la question fut la suivante : le manque de résultats vient-il des données (manque de signal) ou de l'algorithme codé (manque de savoir-faire) ?

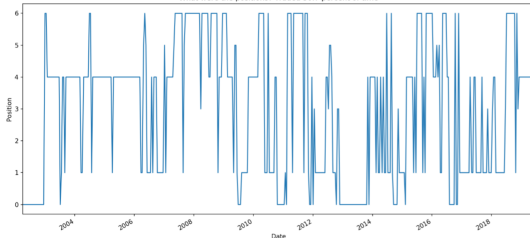
—→ Retour à une librairie complète aux algorithmes éprouvés : **Open AI Gym** (nombreux algorithmes, performance calculatoire, intégration de TensorBoard, exceptionnelle modularité) : environnements standards de RL

Nouveaux résultats

What was the last episode like?



What were the positions? Traded 30.7 percent of time



Des résultats encourageants

- Une fréquence de trading raisonnable ;
- Un indice de référence battu de beaucoup dans les *backtests* ;
- Une courbe de progression monotone et stable.

Des résultats qui encouragent à entrer dans une phase d'optimisation.

- Des résultats obtenus à l'aveugle sans logique sur les paramètres choisis : optimisation plus efficace qu'un grid-search (exemple : algorithme génétique) ;
- Documentation complète pour transmettre à l'équipe de recherche les fruits de mon travail ;
- Explorer le *feature-engineering* pour créer des variables offrant un signal de trading plus clair ;
- Explorer l'Inverse Reinforcement Learning (IRL) pour faire face à la question délicate de la fonction de récompense.

Des questions ?