

Gestion de portfolio par un RL *trading* BOT

Etape 2 : Résultats intermédiaires

Projet - Planification en Intelligence Artificielle IFT 602 / IFT 708

Présenté par :
Tahar Adam AMAIRI
Victor CHAU
Omar CHIDA
Céline ZHANG

Professeur : Froduald Kabanza
Assistants : D'Jeff Nkashama & Jordan Félicien Masakuna



Plan de présentation

- I. Introduction
- II. Etat de l'art
- III. Méthodologie
- IV. Démonstration



I. Introduction

- Précédemment à l'étape 1 :
 - **Sujet** : Trading bot utilisant du reinforcement learning
 - **Source** : *An application of deep reinforcement learning to algorithmic trading*. Thibaut Théate, Damien Ernst
 - **Tâches** : Visualisation dynamique, trade en temps réel, amélioration des plots
 - **Améliorations** : Essayer d'autres modèles tels que le Proximal Policy Optimization (PPO) à la place du TDQN
- Aujourd'hui à l'étape 2 :
 - Implémentation du PPO
 - Multi-stock trading
 - Multi-data trading



II. Etat de l'art

Proximal Policy Optimization (PPO)

- une méthode de *policy gradient*
- apprend des actions plutôt que des valeurs (ex : Q-values)
- Développé en réponse au TRPO



II. Etat de l'art

Proximal Policy Optimization pour le trading

- Données
- Espace d'action
- Fonction de récompenses
- Trading en temps réel



II. Etat de l'art

Proximal Policy Optimization vs TDQN

- On-Policy vs. Off-Policy:
- Gestion des espaces d'action continue
- Convergence et performance
- Politiques stochastiques



II. Etat de l'art

Proximal Policy Optimization vs TDQN

- Réseau neuronal
- Récompenses différées
- Espaces d'état à haute dimension
- Moins de réglages des hyperparamètres



III. Méthodologie

- Amélioration de l'environnement GYM pour ajouter PPO :
- Actor-Critic Network :
 - **Linéarisation** : extraction des caractéristiques des données
 - **Normalisation** : éviter l'étirement des plages de valeurs
 - **LeakyReLU** : éviter la mort des neurones, réduction de la corrélation
 - **Dropout à 20%** : généralisation au modèle, augmente la robustesse
- ADAM + MSE Loss : ajustement gradient + convergence stable
- Multi-stock trading : gestion parallèle des cours boursiers
- Multi-data trading : apprentissage sur les cours de crypto-monnaie



IV. Démonstration

Allons voir le code !



Références bibliographiques

- [1] Thibaut Théate, Damien Ernst. *An application of deep reinforcement learning to algorithmic trading* : <https://www.sciencedirect.com/science/article/pii/S0957417421000737>
- [2] Code source du *trading* l'algorithme de Thibaut Théate et Damien Ernst : <https://github.com/ThibautTheate/An-Application-of-Deep-Reinforcement-Learning-to-Algorithmic-Trading>
- [3] Hongyang Yang, Xiao-Yang Liu, Shan Zhong, Anwar Walid. Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3690996
- [4] Wouter van Heeswijk. Proximal Policy Optimization (PPO) : <https://towardsdatascience.com/proximal-policy-optimization-ppo-explained-abed1952457b>