

# Deep Learning in Finance

## TP3 : GAN

5 novembre 2024

### Buts

- Implémenter un GAN pour les séries temporelles financières
- Comprendre comment juger les résultats obtenus
- Comprendre et résoudre le problème des réseaux denses
- Construire des GANs primitifs à améliorer la semaine prochaine.

### Données financières

Utiliser le fichier de données mis à disposition sur edunao, ou télécharger des données financières et calculer des rendements ajustés.

## 1 Réseaux denses naïfs

Utiliser le squelette de notebook fourni sur Edunao.

1. Définir un objet generator d'entrée à  $N_F$  dimensions et de sortie à  $T$  dimensions, profond et dense. Notez que  $N_F > T$  est nécessaire a priori.
2. Définir un objet discriminator d'entrée à  $T$  dimensions et de sortie à 1 dimensions, profond et dense, avec activation finale adaptée au problème.
3. Définir un objet **gan** (voir notebook).
4. Entraîner cet objet **gan** avec des données financières de sorte que les fonctions de coût du discriminateur et generateur soient relativement proches. Il faut éviter que la fonction de coût du discriminateur tende vers zéro. Si c'est le cas, utiliser un diviseur plus grand pour la vitesse d'apprentissage. En cas de nécessité, arrêter l'entraînement pour des valeurs de fonctions de coût acceptables.
5. Pour 3 vecteurs d'entrée du générateur, tracer les séries temporelles des rendements ou de leur somme cumulée. Obtenez-vous des séries très différentes ?
6. Test des faits stylisés : générer 1000 séries artificielles.

- (a) Est-ce que  $|r_t|$  a des queues grasses ? (tracer la version `flatten()` de la matrice générée en fonction du temps ; tracer  $P_{>}(|r|)$  avec des axes lin-log (voir cours))
- (b) Tester si la fonction d'autocorrélation de  $r_t$  est essentiellement nulle.
- (c) Tester si la fonction d'autocorrélation de  $|r_t|$  ne décroît pas trop vite.

## 2 Réseaux denses moins naïfs

- 1. Définir le générateur comme précédemment, mais en ne permettant pas de biais dans les poids.
- 2. Refaire les points 1.1 à 1.6. En particulier au point 5, est-ce que les séries temporelles sont très différentes ?

## 3 Faits stylisés

- 1. Est-ce que les rendements artificiels ont les mêmes moyenne et variance que les rendements réels ?
- 2. Quelle est l'influence des fonctions d'activation du générateur sur sa capacité à reproduire des rendements à queue grasse ? Si vous deviez inventer une fonction d'activation qui facilite la génération de rendements de grande intensité, quelle serait-elle ?
- 3. Est-ce que ces réseaux génèrent des séries temporelles avec une volatilité à longue mémoire ?
- 4. Pour les données réelles, tracer en fonction du temps la probabilité estimée par le discriminateur que les données soient réelles.