

Python TP9 - Tensorflow

Machine learning

Certains d'entre vous ont pratiqué avec PyTorch, ici nous allons nous servir de la bibliothèque **Tensorflow** (<https://www.tensorflow.org/>) qui est l'autre alternative pour développer des programmes de machine learning.

Nous utiliserons principalement **keras** qui était une bibliothèque indépendante et qui est désormais intégrée dans tensorflow (depuis la v2).

Ceux qui n'ont pas suivi la filière IA en 4A auront, en plus de keras, à se renseigner sur les réseaux de neurones (CNN et LSTM).

Séries temporelles

Nous allons appliquer dans ce TP du *machine learning* pour prévoir les valeurs futures d'une série temporelle. Une série pour laquelle nous avons un historique important, valeurs toutes les minutes sur une vingtaine d'années, est la valeur de l'euro en dollar américain appelé **Forex EUR/USD** (*FOREign Exchange*).

Le plus difficile et le plus intéressant est de prévoir si le cours va monter ou descendre (**sens** ou signe) dans un certain temps appelé horizon (par exemple h=5 minutes). L'autre information est l'amplitude de cette hausse ou baisse, on ne s'y intéressera pas ici. Dans ce problème, un taux de 50 % de réussite (bonne prédiction du sens) est nul, 55 % est moyen et 60 % est bon.

Les données fournies (DAT_XLSX_EURUSD_M1_2018.xlsx et DAT_XLSX_EURUSD_M1_2019.xlsx) proviennent du site <http://www.histdata.com/>. La première colonne contient la date-horaire, les 4 suivantes le open, high, low, close correspondant aux cours en début de minute, max, min et fin de minute. Il y a une dernière colonne qui généralement correspond au volume échangé ou au nombre de changements du cours pendant la minute mais elle n'est pas renseignée ici.

Les 2 fichiers contiennent un très grand nombre de data (plus de 372000 chaque). A vous de sélectionner les données utiles pour l'apprentissage, la validation et les tests.

Travail du TP

1. Lire les données dans le fichier (colonne E **close** qui correspond à la dernière valeur de l'intervalle ici 1 minute) et afficher la courbe.
2. Mettre en place une évaluation **copie carbone** qui considère que le cours monte s'il vient de monter et réciproquement qu'il va descendre s'il vient de descendre pendant la dernière minute.
3. Utiliser keras pour mettre en œuvre un **CNN 1-D** avec une fenêtre de valeurs passées (taille à choisir, par exemple 120 minutes).
4. Utiliser keras pour mettre en œuvre un **LSTM**.
5. Vous avez fini les TP proposés alors faites un programme de votre choix qui vous fera découvrir de nouvelles notions liées à python (une nouvelle bibliothèque, un petit jeu ...)