

Stage : Apprentissage machine supervisé et non supervisé, ENSTA, 20 et 21 octobre 2020

Effectif : environ 20

Auteur du CR : Colas Bardavid

/ Contexte

Le stage a eu en distanciel pendant les vacances de la Toussaint, les cours ont eu lieu sur Zoom ; les formateurs s'appuyaient sur des supports de cours qui nous ont été donnés. Les TP, très soigneusement préparés, interactifs, utilisaient la plateforme colab.research.google.com.

Les participants au stage étaient des professeurs de toutes disciplines, avec des niveaux variés en informatique. Les formateurs ont pris en compte cette contrainte pour que le stage soit profitable à tous.

L'ambiance était studieuse et agréable ; le stage se déroulant en distanciel, nous n'avons que peu eu l'occasion d'avoir des échanges plus informels (entre collègues ou avec les formateurs).

/ Déroulé du stage, première journée

Elle a été animée par Zacharie Alès.

Le matin, nous avons appris les techniques d'analyse en composante principale, qui permettent de détecter dans un nuage de points de \mathbb{R}^N quels sont les « attributs » les plus discriminants. Je ne connaissais pas cette technique simple qui utilise un peu d'algèbre linéaire et bilinéaire mais dont la puissance est surprenante. En application, nous avons codé des algorithmes de compression d'images.

L'après-midi, nous avons étudié les techniques de partitionnement d'ensembles de données. Nous avons entre autres vu :

- les techniques hiérarchiques ;
- la méthode dite des « k-means ».

En TP, nous avons expérimenté l'efficacité de ces techniques en utilisant le package scikit-learn où ces algorithmes sont déjà implémentés.

/ Déroulé du stage, deuxième journée

Elle a été animée par David Filliat et concernait les techniques d'apprentissage supervisé :

- les « Support Vector Machine » (SVM) ;
- les forêts aléatoires d'arbres de décision ;
- les réseaux de neurones.

En TP, nous avons expérimenté l'efficacité de ces techniques en utilisant le package scikit-learn où ces algorithmes sont déjà implémentés.

/ Conclusion et remerciements

Si je ne devais retenir qu'une seule chose de ce stage, ce serait la technique d'analyse en composantes principales et ses applications à la compression d'images qui m'ont beaucoup enthousiasmé !

J'ai aussi apprécié de voir les outils utilisés en TP et la façon dont les TP étaient organisés. J'ai d'ores et déjà changé des éléments de mes propres TP en m'inspirant des pratiques pédagogiques mises en place.

La présentation des techniques d'apprentissage par des réseaux de neurones m'a permis de mieux comprendre les enjeux et d'en avoir une compréhension plus précise.

Ce stage me permettra d'avoir plus d'assurance dans la préparation d'un cours sur l'une ou l'autre de ces techniques ; je pense en particulier à la méthode dite des «k-means » qui figure dans les (projets de) nouveaux programmes d'informatique commune. Les TP dématérialisés sont à ce titre une ressource très précieuse.

Nous avons beaucoup utilisé des algorithmes déjà codés dans scikit-learn et j'aurais bien aimé que l'on code nos propres versions de ces algorithmes.

Je souhaite remercier Zacharie Alès et David Filliat pour la qualité de leurs cours et de leurs TP ainsi que leur investissement et leur disponibilité.