
Projet de Machine Learning Avancé (MLA)

Master en Ingénierie des Systèmes Intelligents & Système Avancées et Robotique 2023-2024

1 Présentation du projet

L'objectif de ce projet est de mettre en pratique, au travers d'un problème concret, les concepts théoriques et méthodologiques d'apprentissage profond vus en cours. Un aspect particulièrement important en recherche et en ingénierie est sa reproductibilité, c'est-à-dire de ré-implementer en détails et à l'identique un algorithme et de reproduire l'ensemble des résultats expérimentaux obtenus. **Votre mission sera de reproduire les résultats expérimentaux obtenus dans un article de recherche fondé sur des réseaux de neurones profonds.**

2 Modalités pratiques

Le projet est à réaliser par **groupe de 4 étudiants**. Il consiste à implémenter en TensorFlow ou pyTorch en totalité l'architecture neuronale proposée dans un article de recherche et la reproduction de tout ou partie des résultats expérimentaux qui y sont présentés. L'article sur lequel vous choisirez de mener votre projet est l'un des 9 articles listés sur le Moodle.

2.1 Sources externes et plagiat

Aujourd'hui la recherche est de plus en plus ouverte, et il est désormais fréquent qu'il existe une ou plusieurs implémentations d'un algorithme présenté dans un article. Si il n'est pas interdit de lire ces implémentations et parfois de s'en inspirer, vous devez produire une implémentation personnelle de l'algorithme, dans son intégralité.

Toute tentative de plagiat de sources externes, en partie ou en intégralité, fera l'objet d'une attention particulière et sera sévèrement sanctionnée.

2.2 Utilisation des GPU

Vous avez la possibilité d'utiliser les processeurs GPU du Master. Ils sont en accès libre à tous les étudiants disposant d'un compte. Il est donc primordial de respecter des règles de bonne conduite et de faire preuve d'une utilisation raisonnable pour que chaque groupe bénéficie d'un accès équitable à ces ressources. De plus amples informations sur l'utilisation des GPU vous seront fournies sur Moodle **Tout abus dans l'utilisation des GPU sera sanctionné.**

Vous pouvez par ailleurs utiliser des ressources externes : ordinateurs personnels ou services en ligne tels que google collab (<https://colab.research.google.com/>). Sur cette plateforme par exemple, les temps d'apprentissage sont limités (12h maximum) mais ils peuvent vous permettent de valider votre code avant de l'exécuter sur le serveur du Master. Par ailleurs, il est possible d'utiliser des GPU (par défaut les notebooks s'exécutent sur CPU). Dans le menu **Exécution**

sélectionnez **Modifier le type d'exécution**. Vous pouvez alors choisir un GPU comme accélérateur matériel.

3 Modalités de rendus et formats attendus

Pour réaliser ce projet, vous utiliserez le service de gestion de code GitHub. L'ensemble des projets seront regroupés au sein de l'organisation **@Intelligent-Systems-MSc**

A l'issue du projet vous devrez nous fournir :

- Un **dépôt git** organisé et structuré. Le dépôt devra comprendre **code source organisé et commenté**, un **README** décrivant les instructions nécessaires à l'installation et au lancement du programme d'entraînement et de prédiction (dépendances, options au lancement...), et un notebook de démo permettant d'exécuter l'intégralité du code et des expériences permettant de reproduire les résultats de l'article.
- Un **rapport** de 6 à 8 pages rédigé **en anglais** à partir du modèle LaTeX fourni sur la plateforme Moodle de l'UE, et présentant sous la forme d'un rapport technique de recherche le contexte du projet, la présentation synthétique de la solution ré-implémentée de l'article et des bases de données utilisées en précisant les différences avec l'article de référence (données, architecture, fonctions de coût par exemple) en justifiant les choix qui les ont motivés, la présentation des résultats expérimentaux obtenus et leurs comparaison par rapport à ceux de l'article de référence, ainsi qu'une discussion critique à partir des résultats obtenus.
- Une **vidéo de 5 minutes** de présentation du projet. Il n'y a pas de recommandations strictes à suivre pour cette vidéo, vous pouvez par exemple diffuser une présentation de type Powerpoint avec une incrustation vidéo de votre visage. La seule contrainte est que tous les participants au projet prennent la parole dans la vidéo.

Pour la **partie code**, vous veillerez tout particulièrement :

1. Aux bons usages de programmation : les codes sources doivent être écrit en **langage Python brut**, préférablement en utilisant une implémentation orientée objet définissant les classes et les méthodes spécifiques pour votre architecture ; seule la démo devra faire recours à un notebook pour simplifier l'exécution de la démo ;
2. A la bonne utilisation de votre git : c'est un outil de versionnage de code permettant le travail collaboratif sur un projet informatique. Il doit être **organisé de manière claire et lisible**, et **l'ensemble des étudiants doivent y contribuer régulièrement en fonction de leurs apports réels**. Vous veillerez à respecter ces bons usages en vous inspirant des nombreux tutoriaux ou git existants et librement accessibles en ligne.

4 Barème et dates importantes

Dates importantes :

- **3 Novembre 2023** : lancement des projets
- **5 Novembre 2023** : constitution des groupes
- **Novembre - Début décembre** : suivi d'avancement du projet (voir agenda)
- **15 Décembre 2023 à 23 :59** : date de rendu de projet (rapport, code et vidéo)
- **Semaine du 18 Décembre 2023** : soutenance orale (20mn de questions par groupe)

La note finale sera établie avec le barème suivant :

- **Rapport (/40)**
La note de rapport évalue le projet sur le fond (introduction, présentation de l'algorithme

- (15%), présentation détaillée des spécifications nécessaire à la ré-implémentation (30%), présentation des bases de données utilisées, présentation des résultats expérimentaux obtenus et comparaison par rapport à l'article de référence (40%), et sur la forme (qualité de la rédaction, mise en page, clarté du discours (15%)) ;
- **Code (/30)**
La note prendra en compte l'organisation et la clarté de dépôt Git, la contribution équilibrée de l'ensemble des membres de l'équipe, et la présence d'un code de démonstration permettant de reproduire les principaux résultats exposés dans le rapport ;
 - **Soutenance (/30)**
La note portera sur la pédagogie et la clarté du propos (40%) tout autant que sur qualité des la vidéo (30%), et la pertinence des réponses aux questions (20%)

La note prendra en compte l'investissement, le sérieux, la régularité, et la pertinence du travail fourni tout au long du projet.

Le travail est à effectuer en groupe mais **les notes sont individuelles**. La note finale sera donc le reflet du travail produit collectivement ainsi que de l'investissement de chaque membre de l'équipe (contributions dans la production du code sources, suivi, soutenance)