

# 2025 3A IABD Rattrapage Projet Annuel

## Résumé :

Les étudiants devront réaliser en individuel une version simplifiée du projet annuel et le présenter lors d'une soutenance de rattrapage. Les simplifications sont indiquées par les ratures et **ajouts** ci-dessous.

Les soutenances auront lieu de Jeudi 11 Septembre 2025 à partir de 18h.

## Cadre général

Le résultat du projet annuel sera d'une part un rapport d'un minimum de **10 20** pages présentant l'étude des performances de tous les algorithmes et modèles sur la problématique choisie.

Les étudiants seront fortement encouragés à utiliser l'outil Matplotlib ou Tensorboard pour produire les courbes et graphes retraçant leurs expérimentations ainsi que pour réaliser un document aisément interactif. Une attention toute particulière sera à apporter à l'étude de l'impact des différents paramètres des algorithmes étudiés sur la rapidité de convergence de ceux-ci.

D'autre part, les étudiants devront remettre une application (sous la forme d'un client lourd, d'un site web ou d'une application mobile) interagissant avec une API hébergeant différents modèles pré-entraînés grâce aux datasets appropriés et préalablement constitués.

L'essentiel des démarches des étudiants, des résultats obtenus et de leurs analyses devront être présentés également lors d'une soutenance.

Avant d'appliquer les algorithmes et modèles vus en cours à la problématique choisie, il sera impératif de démontrer la justesse de l'implémentation de ces derniers sur les cas de tests proposés. Ainsi, il est proposé d'implémenter ceux-ci sur des jeux de données classiques (données linéairement séparables, non linéairement séparables, tâches de classification, tâches de régression, etc.) tels que vus en cours de Machine Learning.

Les modèles et algorithmes à appliquer au projet de test sont :

- ~~Modèle linéaire~~

- Perceptron Multi Couches
- ~~Radial Basis Function Network~~
- SVM ou proposition du groupe d'étudiants

L'ensemble de ces modèles et algorithmes devront être implémentés en C, C++, Rust, **OU PYTHON** de manière à pouvoir être aisément utilisés comme une bibliothèque dynamique manipulée depuis des scripts python (optionnellement également en Unity).

Ces mêmes modèles et algorithmes devront être appliqués à la recherche d'un modèle tentant de solutionner une problématique applicative complexe, pour laquelle une implémentation humaine serait difficile (ex. distinction entre deux classes d'images).

Devront être mis en évidence par le biais de courbes les phénomènes vus en cours tels que (liste non exhaustive) :

- Le sous apprentissage
- Le sur apprentissage
- La mise en évidence de biais dans la base d'exemples
- La difficulté de trouver le modèle « suffisamment complexe pour correctement traiter les données d'apprentissage » mais « suffisamment simple pour bien généraliser »
- Etc.

Il est bon de noter que l'implémentation sera nécessaire pour obtenir des résultats, mais qu'il ne s'agit que de la première étape. L'intérêt principal du projet se situant dans la capacité des étudiants à commenter les résultats obtenus et à porter un regard critique sur l'ensemble des outils vus au travers de différents cours et à l'utilité de chacun face à une problématique issue du monde réel.

Les étudiants seront également encouragés à proposer d'autres modèles et algorithmes si ceux-ci s'avèrent pertinents face au choix de leur problématique. De même ils pourront tout à fait faire usage de la bibliothèque fournie de tensorflow/keras à titre de comparaison de leurs implémentations personnelles.

L'utilisation de GIT sera ~~obligatoire~~ conseillée pour assurer un suivi du projet aisé pour l'ensemble des membres du groupe ainsi que de l'encadrant.

L'utilisation d'implémentations externes des algorithmes et modèles vus en cours est proscrite (hormis à titre comparatif).

Le travail de chaque membre du groupe devra être clairement identifié (header de fichier, document de suivi, etc.)

Pour entraîner leur modèle dans le but de l'utiliser au sein d'une application cliente, les étudiants devront se munir d'un dataset approprié.

~~Plusieurs applications sont proposées, mais si les étudiants souhaitent soumettre une autre problématique et qu'ils disposent d'un moyen de constitution d'un dataset pertinent, il pourra éventuellement être validé par l'enseignant.~~

Les couples problématiques/datasets envisagés sont les suivants :

Classification :

- Application permettant de catégoriser une image (photo) soit en classe 'violon', soit en classe 'contrebasse', soit en classe 'harpe'
- Application permettant de catégoriser une image (photo) soit en classe 'écureuil', soit en classe 'hibou', soit en classe 'renard'
- Application permettant de catégoriser une image (photo) soit en classe 'fenêtre', soit en classe 'porte', soit en classe 'portail'
- ~~- Application permettant de catégoriser une image (photo) soit en classe 'Chat' soit en classe 'Chien' soit 'autre'~~
- ~~- Application permettant de différencier une image (screenshot) issue d'un RTS, d'un MOBA ou d'un FPS~~
- ~~- Application permettant de différencier des lettres manuscrites~~
- ~~- Application permettant de différencier une photo d'un match de foot d'un match de rugby~~

~~—Application permettant de différencier les drapeaux (photos) issus de différents pays~~

~~—Application permettant de différencier le genre d'un film en fonction de son affiche publicitaire~~

Régression :

~~—Application permettant de prédire l'âge d'une personne à partir d'une photo de son visage (attention à la constitution du dataset)~~

~~—Application permettant de prédire les calories d'un plat/produit en fonction de sa composition~~

~~—Application permettant de prédire la qualité d'une position de jeu aux Echecs~~

~~—Application permettant de prédire la qualité d'une position de jeu aux Dames~~

~~—Application permettant de prédire la note (imdb) d'un film en fonction de différents paramètres accessibles (attention au texte !)~~

Un soin tout particulier sera apporté lors de la présentation des travaux réalisés et à l'étude des hypers paramètres du PMC.

Le code produit devra être parfaitement maîtrisé, toute impossibilité d'explication de ce dernier pendant la soutenance sera sanctionnée.

*Bon courage à tous !*