

Fondamentaux de l'Apprentissage Automatique - 8INF867

UQAC – Été 2024

Travail d'équipe

Date de remise du travail : 26/06/2024

Modalité de dépôt : Espace de dépôt sur Moodle

Documents à rendre : Code source + rapport des résultats

Evaluation et Pondération : Barème : sur 100, Pondération : 0.7

Notes importantes :

- Il faut absolument commenter le code de chaque question et interpréter le résultat.
- Attention au plagiat !

Projet d'équipe : Date de remise le 26-06-2024

Enoncé projet d'équipe

Espace de dépôt projet d'équipe

Liste des équipes :

https://docs.google.com/document/d/1JvryKO_i-VfUIW2y13V6fceKRsLHdD3TVWw

Le travail demandé dans ce projet d'équipe s'articule autour de la création d'un modèle de classification supervisée ou non supervisée des données de votre choix. Les données doivent être de type **image**, **son** ou **texte**. Vous pouvez trouver des données sur des sites comme :

<https://www.kaggle.com/>

<https://github.com/>

Exemples d'applications :

- Identification d'un email comme spam ou non
- Identification d'objets sur des images
- Reconnaissance faciale / d'objets, ...
- Segmentation d'image
- Convertir un Texte en Speech ou l'inverse
- Etc.

Par contre, votre projet doit assurer tous les points suivants :

1- (Sur 20 points) Préparation des données : les tâches de cette étape peuvent varier d'un projet à un autre selon le data set choisi. Voici quelques exemples de pré-traitements de données :

- Traitement des valeurs manquantes
- Traitement des valeurs aberrantes
- Formatage des données
- Augmentation des données
- Normalisation des données
- Numérisation des colonnes
- Traduction des données, etc.

Important : faites l'effort de trouver un data set intéressant à pré-traiter (c'est-à-dire pas très clean).

2- (Sur 15 points) Réduction de dimension.

3- (Sur 35 points) Utilisation de trois modèles de **de votre choix**. Vous pouvez utiliser des modèles de Machine Learning ou de Deep Learning. Si la tâche de vos modèles est la classification supervisée, préciser la stratégie de validation utilisée (k-fold cross validation, training/test split, etc.) et justifier ce choix.

4- (Sur 15 points) Evaluation des modèles via les métriques de validation adéquates.

5- (Sur 10 points) Interface utilisateur qui permet de :

- Téléverser ou de saisir un nouvel objet à tester.
- Afficher les résultats des données par chacun des 3 modèles.
- Afficher les métriques d'évaluation de chaque modèle.

6- (Sur 5 points) Evaluation par les pairs : chaque membre d'une équipe doit évaluer (sur 5 points) de manière juste tous les membres de son équipe en se basant sur ses efforts et son engagement dans la réalisation du projet. Chaque membre d'une équipe va individuellement rendre les notes qu'il attribue à ces coéquipiers. Faites une bonne répartition et allocation des tâches dès le départ pour éviter toute sorte de conflit.

Bon travail