**Exercice semaine 9 - data science**

**Date limite de soumission des réponses: *mercredi 24 mai 2023 à 12h***

**Discussion en ligne/présentiel : *Vendredi 26 mai 2023 de 18h à 20h***

**Thématiques abordées : Deep learning et mise en production**

**Instructions :** Utilisez Python pour traiter les questions. Aucune autre application logicielle n’est acceptée pour l’analyse des données. ***Il est souhaitable de fournir votre éditeur de codes annoté avec votre devoir.***

Le travail de cette semaine porte sur la classification d’images. Il est question de prédire si l’image fournie est un chat ou un chien (avec éventuellement les probabilités correspondantes). La classification d’images fait partie du domaine de la vision par ordinateur et utilise un sous-domaine de réseau de neurones artificiels appelé **CNN (convolutional neural network ou réseau de neurone convolutifs).**

Les données sont disponibles à l’adresse URL : <https://drive.google.com/drive/folders/1R--8A6Sr1WReld0zUPrm2XmjuysiMmNS?usp=sharing>

Les données sont contenues dans 4 dossiers compressés (à décompresser pour obtenir les images individuelles). Les dossiers nommés train\_cat\_1 et train\_cat\_2 contiennent les images de chat et ceux nommées train\_dog\_1 et train\_dog\_2 contiennent les images de chien.

**Partie 1 – Pré-traitement des images (50 pts)**

1. Importer et montrer 10 images de chat et 10 images de chien dans votre environnement de travail.
2. Quelles sont les valeurs minimales et maximales de pixel des images importées à la question 1.
3. Afficher une image de chien (chien\_noir) et une image de chat (chat\_noir) en noir et blanc.
4. Quelles sont les dimensions de l’image cat.0 et de l’image dog.0 ?
5. Quelles sont les dimensions de l’image chien\_noir et chat\_noir ?

**Partie 2 – Développement de modèles de prédiction (100 pts)**

**Nous allons utiliser tensorflow et Keras pour le développement du modèle de classification de chat et de chien à partir des données fournies dans ce projet.**

1. Expliquer succinctement le fonctionnement d’un CNN.
2. Développer et valider un modèle de prédiction de chat et de chien en indiquant :

* La procédure de mapage des étiquettes (chat ou chien),
* La procédure de création des données d’apprentissage et de validation,
* Le modèle de CNN utilisé.

1. Encapsuler le modèle créé dans un fichier.
2. Tester les performances de votre modèle sur les données non vues par la machine disponibles à l’adresse URL :

<https://drive.google.com/file/d/1WTXkrchaXXSmK1Df2dbq7hLUEwWBL6yO/view?usp=sharing>

**N.B : Vous devez créer un jeu de données test avec les étiquettes correspondantes.**

**Partie 3 – Déploiement du modèle (50 pts)**

1. Déployer votre modèle pour une utilisation par le public.

**The right way to learn data science is to do data science**

**Good luck**