



## Plan de Travail Prévisionnel

Dans ce livrable se trouve le plan de notre futur travail sur la classification d'image.

### 1. Thème choisi

- ✓ Classification des différentes races de Chien

### 2. Algorithme

- ✓ Visual Transformers

### 3. Raisons du Choix de cet algorithme :

- ✓ Meilleure Précision en général que les algorithmes de CNN sur les problèmes de Classification (Cifar10...)
- ✓ Division d'images en Patches
- ✓ Disponible en modèles pré entraînés

### 4. Sources Identifiées

- ✓ Vision Transformers provably learn spatial structure:  
<https://arxiv.org/abs/2210.09221>
- ✓ Understanding the Vision Transformer and Counting Its Parameters  
<https://medium.com/analytics-vidhya/understanding-the-vision-transformer-and-counting-its-parameters-988a4ea2b8f3>
- ✓ Image classification with Vision Transformer  
[https://keras.io/examples/vision/image\\_classification\\_with\\_vision\\_transformer/](https://keras.io/examples/vision/image_classification_with_vision_transformer/)



- ✓ Visual Transformers Documentation : [https://huggingface.co/docs/transformers/model\\_doc/vit](https://huggingface.co/docs/transformers/model_doc/vit)
- ✓ Fine-Tune Visual Transformers for Image Classification with Transformers : <https://huggingface.co/blog/fine-tune-vit>
- ✓ Vision Transformer (ViT) Fine-Tuning : <https://www.kaggle.com/code/raufmomin/vision-transformer-vit-fine-tuning>

## 5. Dataset

- ✓ Stanford Dog Dataset

Nous avons choisi la jeu de données du projet P6 qui est le Stanford Dog Dataset à l'adresse suivant :

<http://vision.stanford.edu/aditya86/ImageNetDogs/>

## 6. Méthode d'évaluation

- ✓ Evaluation de la Précision et perte des modèles  
Nous utilisons la méthode `model.evaluate` pour évaluer la performance du ViT et la comparer a celui du Vgg16 jugé comme meilleur modèle dans le Projet P6

# Projet P7 Opendclassrooms : Développez une preuve de concept

