

# Projet de Machine Learning : Consignes Générales

## Objectif Général

L'objectif principal de ce projet est d'appliquer et d'approfondir les concepts de Machine Learning et de Deep Learning vus au cours, en travaillant en **groupe de 3 étudiants**. Vous devrez choisir un sujet parmi les 3 catégories proposées (Classification d'image Avancée, ). Ces projets vous demandent de rechercher ou utiliser les données adéquates, prétraiter ces données, concevoir, entraîner, évaluer et analyser plusieurs modèles à l'aide de TensorFlow et Keras. Une part importante du projet consistera à **expérimenter** différentes approches et à **documenter rigoureusement** votre démarche.

## Inscription du Groupe et Choix du Sujet

Avant de commencer le développement, chaque groupe doit **impérativement s'inscrire** en indiquant :

1. Les **Noms et Prénoms** de ses 3 membres.
2. L'**énoncé exact du projet choisi**. Donc, la **catégorie** du projet ainsi que le **dataset** que vous avez choisi. Exemples : "Catégorie 1 Chien vs Chat" ou "Catégorie 2 SMS spam"

Cette inscription se fait via le fichier partagé suivant : **Feuille d'Inscription Projet ML**

Veuillez remplir ces informations dès que possible après la formation de votre groupe et le choix de votre sujet, vous avez jusqu'au 25/04/2025. Après cette date, les groupes ne sont plus modifiables. Il est impératif de s'inscrire pour pouvoir passer l'examen.

## Exigences Générales

1. **Soumission du Projet** :
  - Le projet complet doit être soumis sous la forme d'une **archive ZIP**. Vous devez impérativement utiliser les **Notebook Jupyter** pour le code python.
  - Le nom de l'archive ZIP doit être formé des **prénom et noms des 3 membres du groupe**, séparés par des tirets (par exemple : JeanDupont\_SophieMartin\_PierreDurand.zip).
  - L'archive ZIP doit être envoyée par email à l'adresse `a.tillieux@helmo.be` au plus tard **5 jours avant la date fixée pour votre examen oral**.
  - **Confirmation de Réception** : Vous devriez recevoir un **accusé de réception dans les 24 heures** suivant votre envoi. Si ce n'est pas le cas, vérifiez votre dossier de courriers indésirables (spam), puis **renvoyez votre email** pour vous assurer de sa bonne réception. Ne pas avoir d'accusé de réception signifie potentiellement que votre travail n'a pas été reçu.
  - Le notebook doit être **auto-suffisant** : il doit contenir tout le code (Python, TensorFlow/Keras), les cellules de texte explicatives (Markdown), les visualisations, et la documentation détaillée de vos expérimentations.
2. **Gestion des Données** :
  - **AUCUN jeu de données volumineux ne doit être inclus dans l'archive ZIP.**

- Le notebook **doit contenir le code nécessaire pour télécharger automatiquement les données** depuis une source en ligne accessible (ex: `tf.keras.datasets`, `tensorflow_datasets`, API Kaggle, `wget/curl`). Des instructions claires doivent être fournies si une manipulation minimale est inévitable dans le `README.md`.
  - *Exception* : Pour la catégorie “Classification d’Images”, les quelques images de test prises par les étudiants eux-mêmes peuvent être incluses dans l’archive ZIP (idéalement dans un sous-dossier).
3. **Code et Environnement** :
- Utilisez Python avec TensorFlow et Keras comme framework principal. NumPy, Pandas, Matplotlib sont également attendus.
  - Listez clairement les **dépendances externes** (avec versions) au début du notebook ou dans une cellule dédiée.
  - Un fichier `requirements.txt` doit-être présent dans le projet afin de pouvoir installer les dépendances python de votre projet.
  - Le code doit être propre, commenté pour expliquer les parties non triviales.
4. **Travail en Groupe** :
- Travail d’équipe où la collaboration est attendue. Chaque membre doit avoir une compréhension complète de l’ensemble du projet.

## ***Tâche Centrale : Expérimentation Rigoureuse et Documentation (Évaluation Clé)***

Une part **essentielle** de l’évaluation portera sur **votre capacité à expérimenter de manière méthodique et à documenter vos essais** directement dans le notebook.

Pour chaque étape significative (choix d’architecture, hyperparamètres, prétraitement, etc.), vous devez intégrer :

1. **L’Hypothèse/Justification** : *Pourquoi* cet essai ? Qu’attendiez-vous ?
2. **La Méthodologie** : *Comment* avez-vous procédé ?
3. **Les Observations** : *Quels* résultats (quantitatifs/qualitatifs) ? (Courbes, métriques, tableaux...).
4. **Les Conclusions et Décisions** : *Qu’avez-vous appris* ? Comment cela influence la suite ?

Une **section de synthèse** récapitulant les expérimentations clés, les choix finaux et leur justification est requise.

## ***Auto-Apprentissage Guidé***

Attendez-vous à devoir explorer des concepts non couverts en détail en cours. L’énoncé spécifique de votre projet contiendra des **pistes et mots-clés** pour orienter vos recherches (techniques de prétraitement, couches spécifiques, métriques, etc.). L’autonomie dans cet apprentissage sera valorisée.

## ***Contraintes Techniques***

Votre modèle final doit pouvoir être **entraîné en 30-40 minutes maximum sur un ordinateur portable standard** (CPU récent, sans GPU dédié). Optimisez via :

La **taille/complexité** du modèle.

Le **prétraitement efficace des données en amont** (justifiez vos choix pour limiter la charge d'entraînement).

## ***Modalités de l'Examen Oral***

L'évaluation finale se fera lors d'un **examen oral de 30 minutes**.

**Présence Obligatoire** : Les **3 membres** du groupe doivent être présents.

**Structure de l'Oral (30 min)** :

- **Présentation du Projet (15 minutes max)** :
  - Préparez une présentation synthétique (slides ou via le notebook).
  - Contenu : Objectif, dataset, prétraitement clé, **synthèse de la démarche expérimentale (choix, résultats, décisions)**, modèle final et performances, succès/difficultés, **concepts appris en autonomie**, pistes d'amélioration.
- **Discussion et Questions (15 minutes min)** :
  - Séance interactive avec l'enseignant.
  - **Interrogation Individuelle Possible** : Chaque étudiant peut être interrogé **individuellement**.
  - **Attentes Individuelles** : Chaque étudiant doit maîtriser :
    - Les concepts théoriques du cours et leur application dans le projet.
    - Le code produit par le groupe.
    - L'explication des expérimentations, observations et analyses du notebook.

## ***Critères d'Évaluation Indicatifs***

- Qualité et rigueur de la démarche expérimentale et de sa documentation.
- Pertinence et justification des choix techniques.
- Clarté et propreté du code et du notebook.
- Correction de l'analyse des résultats et des conclusions.
- Qualité de la présentation orale et maîtrise individuelle des concepts.
- Respect des consignes (soumission, format, contraintes, inscription).

Bon travail ! N'hésitez pas à poser des questions et commencez à travailler tôt sur ce projet.