

Projet Machine Learning

Table des matières

I. Conditions du projet et Livrables.....	2
II. Définition Projet ML.....	2
III. Ressources complémentaires pour le projet.....	3
IV. Livrer le tout le Jeudi 05/06/2025 pour 17h30.....	3

Libellé du cours	Durée	Début	Fin
PROJET : Conception et développement d'IA 3J	4.0	lun. 02-06-2025 09:30	jeu. 05-06-2025 17:30

I. Conditions du projet et Livrables

A. Travail Seul ou en groupe : inviter **robinhotton** pour les livrables sur le repo **Github**

B. Livrables :

- Sources en python
- Documentation technique
- Readme.md : notes sur le processus, l'équipe, la répartition du travail, mode d'emploi pour utiliser votre application

II. Définition Projet ML

Objectif : construire une petite application avec Streamlit en python, afin de présenter un pipeline complet de Machine Learning ou Deep Learning.

Ressource : un jeu de donnée vin.csv.

Étapes :

1. **Construction de Streamlit :** Discuter sur l'architecture d'utilisation. Vous pouvez organiser vos pages sous forme de page, sous forme d'onglet, sous forme d'accordéon, en utilisant la sidebar ou non.
2. **Chargement du jeu de données :** Charger vos jeux csv à partir de chemins locaux dans votre application afin de construire vos dataframe. A partir de là, vous devez intégrer des interactions utilisateurs
3. **Bloc de traitement des données :** essayez d'intégrer le maximum de fonctionnalité dans le traitement des données, voici quelques idées intéressantes à implémenter :
 - a. analyse descriptive du dataframe
 - b. graphiques de distribution et pairplot
 - c. corrélation avec la target
 - d. Fréquences
 - e. standardisation
 - f. Gestion des valeurs manquantes
 - g. etc...

Il serait intéressant d'intégrer des interaction utilisateurs, comme par exemple permettre à l'utilisateur de sélectionner les colonnes souhaitées et appliqué sa demande (afficher un graphique, supprimer une colonne, imputer les valeurs manquantes par un choix etc)

4. **Bloc Machine Learning :** essayer d'intégrer un pipeline de traitement avec la possibilité de laisser choisir entre plusieurs algorithmes selon la target détecté. Ensuite, appliquer le split, le fit et les prédictions des données. Vous pouvez permettre à l'utilisateur de prédire sur de nouvelles données, d'enregistrer le model etc

5. **Bloc d'évaluation** : essayer d'intégrer un bloc d'évaluation du model qui vient de tourner et de s'entraîner. Vous pouvez utiliser les métriques ou bien des graphiques.
6. **BONUS : Ajouter des fonctionnalités supplémentaires** : essayer d'ajouter des fonctionnalités pour optimiser l'application comme par exemple, un lazy predict, un gridSearchCV, l'intégration d'un modèle de Deep Learning etc. pas de limite ...

III. Ressources complémentaires pour le projet

A. Ressources E-Learning :

- <https://fr.linkedin.com/learning/python-pour-la-data-science?u=123653769>
- <https://fr.linkedin.com/learning/l-intelligence-artificielle-et-les-systemes-de-recommandation?u=123653769>
- <https://fr.linkedin.com/learning/l-intelligence-artificielle-et-l-apprentissage-par-renforcement?u=123653769>
- <https://fr.linkedin.com/learning/scikit-learn-pour-le-machine-learning?u=123653769>
- <https://fr.linkedin.com/learning/machine-learning-les-algorithmes-d-apprentissage?u=123653769>

B. Ressource Streamlit :

- <https://fr.linkedin.com/learning/python-pour-la-data-science-serie/comment-deployer-rapidement-des-applications-de-data-science-avec-la-librairie-streamlit?resume=false&u=123653769>
- <https://fr.linkedin.com/learning/l-essentiel-d-openai-pour-le-developpement-en-python/demarrer-avec-streamlit-io?resume=false&u=123653769>
- <https://streamlit.io/https://docs.streamlit.io/get-started/tutorials/create-an-app>
- <https://app.dataquest.io/m/900/getting-started-with-streamlit/1/introduction>
- <https://fr.linkedin.com/learning/python-pour-la-data-science-serie/comment-se-creeer-des-opportunites-en-tant-que-data-scientist-maitrisant-le-langage-python?u=123653769>

IV. Livrer le tout le Jeudi 05/06/2025 pour 17h30.