

PROJET FIL ROUGE

PRÉSENTATION DU CAS D'USAGE



Maths et Datascience

Le projet

Contexte :

- Un regroupement de municipalité gèrent la qualité des eaux d'un ensemble de lacs.
- Durant la période de printemps été, le développement de cyanobactéries posent des problèmes de santé publique : impact sur la santé / tourisme / réservoir d'eau potable pour l'agriculture, ...
- Les analyses en laboratoire sont trop longues
- Le regroupement cherche à financer une étude / prototype pour les aider.

Déroulement :

- L'encadrant va jouer le rôle de datascientist / représentant des municipalités
- Le jeu de donnée choisi est un jeu de donnée réel disponible :
- (<https://data.world/us-epa-gov/b908b06e-ec8e-4125-b762-c4ff80f13206>)
- Développé en parallèle d'une étude statistique :
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8168688/>

Le projet

Les objectifs du projet :

- **Effectuer un projet de datascience complet en groupe.**
- **Interagir avec le métier**
- **Comprendre les besoins métiers**
- **Réaliser une(s) solution(s) apportant de la valeur au métier**

Déroulement du projet

- Constituez 4 groupes **maximum**
- Durant chaque séance, entretien de **15mn par groupe** avec l'encadrant qui jouera le rôle du métier.
- Soutenance orale à la dernière séance. Vous avez 20 minutes de présentation, suivis de 5-10 minutes de questions. Les questions seront "technique" (sur les choix techniques posés par un datascientist) et "métier" (posées par le représentant du regroupement de municipalité ne comprenant pas forcément la datascience).

DEADLINES:

- Slides et maquette au plus tard la veille de la date du dernier cours à 23h59. Les premiers à rendre seront les derniers à passer (ce qui est un avantage potentiel)

Les livrables

Les slides de présentation au format pdf.

- Ils doivent expliquer la démarche de l'équipe, justifier les choix et pistes explorée, et démontrer la valeur apportée au métier.
- Chaque personne de l'équipe est responsable d'un use-case (ou piste de l'un) et doit présenter au minima sa slide sur celle-ci.

Une maquette de la solution réalisée.

- Technologie libre.
- Elle doit contenir un fichier readme.md expliquant le cas échéant les fichiers fournis, comment installer ou utiliser la solution et tous les prérequis nécessaires.
- Elle doit être faire l'objet d'une **démonstration lors de la soutenance.**

Nos conseils

Trouver le ou les bonnes applications de la datascience peut être plus déterminant que l'optimisation de votre solution.

- Ne pas foncer "tête baissée" sur le premier use case mais prendre le temps de définir le ou les uses cases les plus **intéressants** (valeur apportée au métier, **faible risque de ne pas marcher d'ici la fin du projet**, ...)
- L'utilisation du machine learning n'est pas obligatoire pour apporter de la valeur au métier.

Le projet est noté sur :

- Les entretiens hebdomadaires avec le client
- La valeur apportée au métier
- La présentation : qualité, clarté, caractère convaincant pour le métier
- La qualité de la démarche de data-science
- La maquette : qualité, facilité d'utilisation par le métier
- Certaines parties du barème pourront être individuelles



SCALIAN

A BRIDGE TO PERFORMANCE

Société SCALIAN

417 L'Occitane - CS77679

31676 Labège Cedex

Tél. : +33 (0)5 61 00 79 79

scalian.com