

PROJET FIL ROUGE

PRÉSENTATION DU CAS D'USAGE



Maths et Datascience

Le projet

Contexte :

- Un e-commerçant vient vous voir.
- Il a des données.
- Il ne sait pas comment tirer de la valeur.
- Il pense que la datascience peut lui permettre de valoriser ses données.

Déroulement :

- L'encadrant va jouer le rôle de l'e-commerçant tout au long du projet.
- Le jeu de donnée choisi est un jeu de donnée réel d'un e-commerçant brésilien (<https://olist.com/>)
- Jeu de donnée disponible librement sur kaggle :
https://www.kaggle.com/olistbr/brazilian-ecommerce?select=olist_order_payments_dataset.csv

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 des groupes de 2.
- Durant chaque séance, entretien de **15mn par groupe** avec l'encadrant qui jouera le rôle :
 - Une fois par semaine : du client
 - Une fois par semaine : du datascientiste (ou du client *si vous en faites la demande*).
- Soutenance orale la dernière séance. Vous avez 20 minutes de présentation + démo, suivis de 5-10 minutes de questions. Les questions seront "techniques" (sur les choix techniques posés par un datascientist) et "métiers" (posées par le e-commerçant ne comprenant pas forcément la datascience).

DEADLINES:

- Slides et maquette au plus tard la veille de la soutenance. Les premiers à rendre seront les derniers à passer (*ce qui est un petit avantage*)

Les livrables

Les slides de présentation **au format pdf.**

- Ils doivent expliquer la compréhension du problème, 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'au moins un use-case** et doit présenter au minima les slides présentant celui-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.**

Des conseils pour le datascientiste

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 barèmes pourront être individuelles si on juge que c'est nécessaire.

Des conseils pour la gestion du projet

Soyez **SMART**

S : Specifié : prenez le temps au début de comprendre les données, le besoin, et spécifier quel besoins exacts vous adressez et les raisons.

M : Mesurable : mesurez la performance de votre solution, et la valeur qu'elle va apporter à l'utilisateur. Si possible, définissez des performances minimums pour qu'elle soit utile au client.

A : Agile : en datascience et ML, il est difficile de prévoir à l'avance les performances, ce qui va marcher et ne pas marcher. Solution : procéder par itérations rapides et/ou incrémentales.

R : Realiste : ne cherchez pas à faire un chatgpt ! Objectif ambitieux mais atteignables.

T : Timez : Mettez des deadlines pour chacune des étapes de votre projet. ***Et pas une unique à 23h59mn59s la veille de la soutenance.***

Last but not least

Une réponse de soutenance : « Toi et Pierre ne nous avez pas dit de faire la même chose »

> C'est voulu !

- Deux interlocuteurs différents, même dans la même société, ne vont pas toujours exprimer le même besoin.
- Lorsque cela arrive, vous devez :
 - Evaluer les différents besoins exprimés
 - En sélectionner un, deux ou un compromis pour des raisons **justifiés**, mesurable et discutées avec les interlocuteurs.
 - Identifiez qui est(sont) décideur(s), et s'assurer que tous sont en accord avec l'arbitrage réalisé.
- En tant qu'ingénieur, il faut être conscient que la réponse « je fais ça parce qu'on m'a dit de le faire » ***n'est pas suffisante***.



SCALIAN

A BRIDGE TO PERFORMANCE

Société SCALIAN

417 L'Occitane - CS77679

31676 Labège Cedex

Tél. : +33 (0)5 61 00 79 79

scalian.com