



Projet Occi

Contexte

L'objectif de ce mini-projet est d'implémenter un algorithme permettant d'étiqueter chaque point d'une trajectoire comme « en mouvement » ou « à l'arrêt ». Le dataset fourni est constitué de trajectoires réelles collectées par notre système dans le magasin pilote.

L'intérêt de cette tâche est qu'elle permet ensuite de décider si un client s'est arrêté ou non dans un rayon pour consulter des produits, et donc d'inférer une certaine mesure du taux de conversion et du niveau d'intérêt client pour ce rayon.

Contraintes techniques

Nous demandons de programmer en **Python** car la majorité de notre code est aujourd'hui dans ce langage, et c'est essentiellement ce langage qu'on te demandera d'utiliser au cours de ton stage.

Tu peux réaliser ce projet à ton rythme, et nous envoyer tes résultats quand tu es satisfait. Il n'est pas censé te prendre beaucoup de temps, le but est surtout d'évaluer ton niveau en algorithmique et implémentation. Si tu bloques ou rencontres un problème particulier, n'hésite surtout pas à nous contacter (louis@occi.io), l'idée est de t'intégrer à une équipe qui travaille ensemble, et de pouvoir t'appuyer sur notre expertise quand tu en as besoin.

Une grande attention sera évidemment portée à la lisibilité et clarté du code fourni, ainsi qu'à la présence de tests du bon fonctionnement du code.

Si tu manques de temps, voici à titre indicatif les points que nous regarderons en priorité. Si tu ne peux pas tout faire, ce n'est pas grave, essaie d'aller aussi loin que possible en un temps raisonnable :

- Charger les trajectoires et les visualiser, individuellement ou plusieurs à la fois.
- Algorithme de détection de point d'arrêt.
- Visualisation des points d'arrêts sur la trajectoire

Dataset

Le fichier zip fournit contient un fichier csv contenant les colonnes suivantes :

- **traj_id** : identifiant d'une trajectoire
- **x** et **y** : les coordonnées du point dans le plan
- **t** : le temps de la mesure
- **error** : l'erreur de mesure de la position, i.e. le rayon du disque (en mètres) dans lequel la vraie position se trouve

Ainsi, pour extraire une trajectoire il faut récupérer les lignes avec l'identifiant de la trajectoire puis trier ces lignes en fonction du temps.