Compte rendu de la réunion Zoom – jeudi 19 novembre

Durée : 40 minutes

**Présents :** Charles Prud’homme, Alexis Schneider, Antonin Bernard, Luana Siquiera, Helder Betiol, Hermine Anceaux

**Données :**

LoudML est utilisé pour de la prévision et pas de la génération pour Charles. Toutefois si on considère qu’une prévisione est nécessaire à chaque injection, cela peut marcher.

Pour savoir si LoudML est approprié pour le projet, il faudrait constituer une petite équipe qui travaille dessus.

Proposition de Charles : Avec la saisonnalité, la tendance et le bruit, on peut générer des polynômes. Ça reprend ce qu’avaient fait les élèves GIPAD. On estime la pente de la montée/descente (tendance), et on décore avec la saisonnalité et le bruit.

On peut ensuite faire une étude comparative avec ce qui a été fait avec LoudML. Le problème de LoudML, c’est que c’est une boîte noire, qu’on ne sait pas vraiment ce qu’il se passe à l’intérieur.

Avec les vrais lapins, certains réagissent exagérément, d’autres beaucoup moins. Le clustering permet de diviser les lapins en différents groupes, puis on peut séléctionner un groupe aléatoirement et générer le tracé adapté.

L’hypothèse d’une distribution en loi normale vient du travail des GIPAD, qui n’ont travaillé que sur la fréquence cardiaque et l’adrénaline. C’est à poursuivre si on peut écarter les autres hypothèses. Il ne faut pas fermer d’approches brutalement.

Cela dit, avec la loi normale, le clustering est plus précis, suit mieux la tendance. Il est probable que les autres répartitions suivent la même loi.

C’est une bonne approche de faire des « mini-experts » par thématique, et tout confronter à la fin.

Pour le clustering, il faut utiliser une bibliothèque qui existe déjà (inutile de tout refaire en moins bien). FastDTW permet de comparer les tendances, mais si on utilise LoudML, est-ce que c’est quelque chose qu’on va utiliser ?

L’objectif est de trouver des outils existant pour les mettre dans une pile logicielle qui répond à nos besoins.

Sous-module seasonal proposé par Charles :

<https://www.statsmodels.org/dev/generated/statsmodels.tsa.seasonal.seasonal_decompose.html>

Tuto :<https://machinelearningmastery.com/decompose-time-series-data-trend-seasonality/>

Il s’agit d’une décomposition additive.

Au niveau de l’importation des données, on trouve parfois des valeurs très élevées pour la fréquence cardiaque. C’est sans doute causé par le lapin qui bouge (bruit).

Pour avoir des valeurs réelles, il vaut mieux aller voir directement Julie.

Quand il y a des valeurs problématiques, on peut imaginer reconstruire les données manquantes. La méthode utilisée jusqu’ici était de supprimer totalement les données aberrantes et de recoller ensemble de chaque côté de l’intervalle supprimé. Peut-être que LoudML peut proposer quelque chose pour faire ça. On risque cependant d’avoir des erreurs dans les tracés.

Quand on a du bruit sur quelques secondes, on peut couper sans problème. Quand il s’agit de beaucoup de secondes, ça devient problématique.

La saisonnalité est plus facile à mimer que la tendance.

**Maquette :**

Charles a contacté le directeur d’études. Il est possible de convenir un rendez-vous en click and collect. Rendez-vous convenu le 20/11 à midi pour récupérer la maquette.

Il faut donner des nouvelles souvent pour ne pas rester bloqués. Il vaut mieux trop solliciter les intervenants que pas assez.