Modélisation du comportement utilisateur dans la publicité mobile

*13/03/2020*

WEEKLY STATUS REPORT

**De** : Joseph AKA BROU, Marwa ELATRACHE, Caroline MARTIN, Tharshika NAGARATNAM, Omar SECK

**À** : Faustine BOUSQUET, El hassan DAOUDI, Samy ZAROUR

**Période** : 02/03/2020 – 13/03/2020

|  |
| --- |
| **Activités réalisées cette semaine** |
| * Fin du nettoyages de données (mois de novembre 2019) → 16go * Fin de l’étape de features engineering * Développement de quelques modèles (Scala, Python, R) :   - Réduction des dimensions (suppressions de variables, SVD)  - Over sampling (Smote)  - Régression logistique, Linear SVM, Random Forrest, Boosting, Réseau neurones |

|  |
| --- |
| **Objectif de la semaine prochaine** |
| * Améliorer nos modèles respectifs * Dégager une méthode. (La plus pertinente d’entre celles développées) * Rencontre avec les experts le 03/02/2020 |

|  |
| --- |
| **Points bloquants demandant une attention particulière des encadrants** |
| * Manque de connaissance sur la pénalisation des faux positifs. (Il va falloir se documenter) |

|  |
| --- |
| **Remarques additionnelles** |
| * Nous sommes désormais tous sur python ou R dans le cadre la recherche de modèle supervisés sur python et R. Utiliser Scala nous a permis de nous familiariser avec le langage. Cependant contrairement à python, Scala n’est pas optimisé pour la recherche de modèles. Scala s’oriente plus pour la prod. Ainsi, nous reviendrons vers scala lorsque nous aurons fixé une solution. |