




Prédiction de flux de marchandises

Mieux piloter son activité
grâce au machine learning



1. Mission

Prédire les volumes de marchandises à livrer

- sur **250 plateformes** de distribution réparties sur l'ensemble du territoire
- à des horizons de temps compris **entre 3 et 6 semaines**

2. Stratégie

- Audit des données historiques disponibles
- Mise en oeuvre d'une approche **machine learning**
 - Fonctionnement en itérations - **3 x 2 semaines**
- Evaluation des résultats en temps-réel en **compétition avec les experts métiers (humains !)**
- Passation aux équipes techniques pour mise en production en autonomie

3. Défis

- Reconstruction d'un historique de 5 ans standardisé prenant en compte :
 - les changements d'outils et de pratiques
 - les évolutions de la structure (fusion de plateformes, ajout/suppression de plateformes)
 - les événements locaux : calendrier scolaire & électoral, prévisions météo
- Construction d'une *baseline** pour évaluer l'algorithme
- Evaluation des marges d'incertitude des prédictions

Cas d'usage par Romain Gauthier

* *baseline* ⇒ prédiction initiale
servant de comparaison

4. Résultats

- **Résultats des prédictions**
 - moins de 15 % d'erreur à 3 semaines
 - moins de 20 % d'erreur à 6 semaines
 - marges d'incertitudes à 95% sur chaque prédiction
 - victoire sur les experts métiers (10 points de mieux en moyenne)
- **Prise en main de l'outil par l'équipe IT du client**
(stack techno open source : Python / XGBoost / Seaborn / Linux)

5. Tips & Apprentissages

- Ne jamais s'engager avant d'avoir eu **TOUTES les données**
corollaire ⇒ ne jamais croire quiconque sur parole quand il s'agit de **données historisées** : il est probable que vos interlocuteurs ne partagent pas votre définition de "*propreté des données*" - et ce n'est absolument pas de leur faute !
- Les données externes peuvent apporter une amélioration des prédictions, mais **l'essentiel se joue sur les données internes** !
- Tout le monde parle Excel ⇒ ne pas hésiter à l'utiliser si nécessaire !

XGBoost & LightGBM sont des bibliothèques extraordinaires (mais ça, on le savait déjà)

Merci !



<https://romaingauthier.io>
bonjour@romaingauthier.io
07 80 97 61 50