



l'école d'ingénierie  
informatique

# TP Composant métier

Compte rendu

Évaluateur : M. MESSAOUDENE

Étudiants : Lucas VARO - Nathan OLIVE - Quentin HASNAOUI - Clément BRIQUET

Terminé ▾

# Sommaire

|                                |   |
|--------------------------------|---|
| Rappel des consignes :.....    | 2 |
| Hypothèse.....                 | 2 |
| Conclusion de L'Hypothèse..... | 3 |

# Rappel des consignes :

- Trouver une problématique
- Émettre au moins une hypothèse et y répondre
- Pas de limite de modèle
- Github : code + fichier d'explication

## Hypothèse

Il est possible de prévoir la réussite d'une inspection sanitaire à une localisation géographique et une temporalité. En partant de cette hypothèse, qu'elles seraient les zones géographiques pour une date donnée qui ont le plus de chance de réussir une inspection, ou à l'inverse d'échouer ?

## Démarche employée

La toute première décision que nous avons eu à prendre consistait à déterminer quelles colonnes de notre ensemble de données nous allions utiliser. Après une analyse approfondie de notre jeu de données, nous sommes parvenus à un consensus : seules les colonnes "**Date\_inspection**", "**APP\_Libelle\_activite**", "**geores**", et "**Synthese\_eval\_sanitaire**" répondaient à nos besoins.

Notre travail c'est ensuite divisé en plusieurs étapes :

- **Analyse** : Dans cette étape, nous appréhendons et étudions notre dataset afin de commencer à déceler ce qu'il est possible de faire avec les données que nous avons, mais aussi les opérations que nous devons réaliser afin de pouvoir efficacement utiliser ces mêmes données (nettoyage, normalisation...). C'est aussi dans cette étape que nous avons dressé nos premières hypothèses.
- **Nettoyage** : Consiste à détecter, corriger et supprimer des erreurs, incohérences ou données inutiles dans un ensemble de données pour assurer la qualité et fiabilité afin de rendre les données exploitables.
- **Choix du modèle / Essai de plusieurs modèles / recherche** : Essai de différents modèles afin de comprendre celui qui correspond le mieux à notre besoin métier, et permet d'obtenir le degré de précision le plus élevé.
- **Apprentissage** : Une fois notre modèle choisi, nous passons à la phase de paramétrage de celui-ci afin de trouver la configuration nous permettant d'obtenir les résultats les plus proches de la réalité possible.
- **Data visualisation et analyse finale** : Data visualisation finale, permettant d'analyser nos résultats prédictifs et de les présenter. C'est là que nous formulons la confirmation ou l'infirmité des hypothèses réalisées en amont.

# Conclusion de L'Hypothèse

En conclusion, en suivant cette démarche méthodique, nous avons pu explorer et analyser les données pour répondre à notre hypothèse initiale. Nous avons obtenu des informations précieuses sur les zones géographiques qui ont plus de chances de réussir ou d'échouer lors d'une inspection sanitaire à une date donnée. Cette analyse peut être utile pour prendre des décisions éclairées en matière d'inspection sanitaire et de gestion des ressources.

Après analyse des résultats de notre modèle, nous avons pu observer qu'il semble possible de prédire la réussite d'un établissement à une inspection, en fonction de ses attributs géographiques (ce qui confirme notre hypothèse initiale). Cependant, notre modèle n'est pour autant pas aussi précis que nous l'aurions souhaité, et afin de pouvoir parfaitement confirmer notre hypothèse, il serait donc encore intéressant de travailler sur celui-ci, jusqu'à obtenir des résultats encore plus précis.