RASSANT Nolan

NDOMBI MOUIRI Caprice

ROUSSEL Benjamin

***RENDU ATELIER DE MACHINE LEARNING***

Dépôt Github : <https://github.com/NolanRassant/atelier_machine_learning>

## Problématique

* Est-ce que l’évaluation (Synthese\_eval\_sanit) diffère selon le type d'établissement (APP\_Libelle\_activite\_etablissement) ?
* Est-ce que les inspections se passent à une certaine période (Date\_inspection) dans l’année ?
* Est-ce que le lieu (Libelle\_commune ou/et Code\_postal) joue sur la note (Synthese\_eval\_sanit) ?

## Hypothèse

Il semblerait que les dates, lieux et types d’établissement peuvent jouer sur la note d'évaluation, et donc le compte rendu de l’inspection.

On constate aussi qu’il y a certains mois pendant lesquels les inspections sont plus fréquentes.

## Réponse

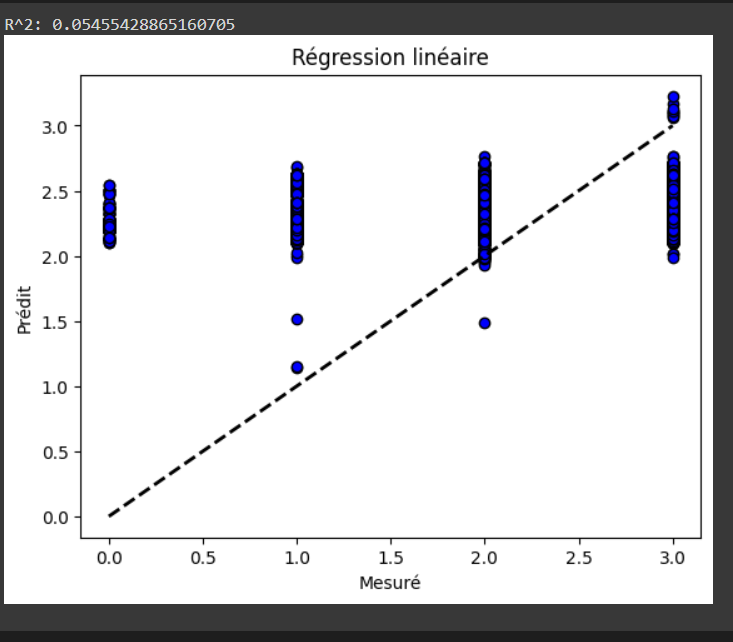
On a fait un modèle de régression logistique avec une précision de 0.54, ce qui veut dire que oui, ses variables jouent en effet sur la note finale et peuvent prédire correctement.

D’après le Random Forest, R² de 0.43, donc pas suffisant

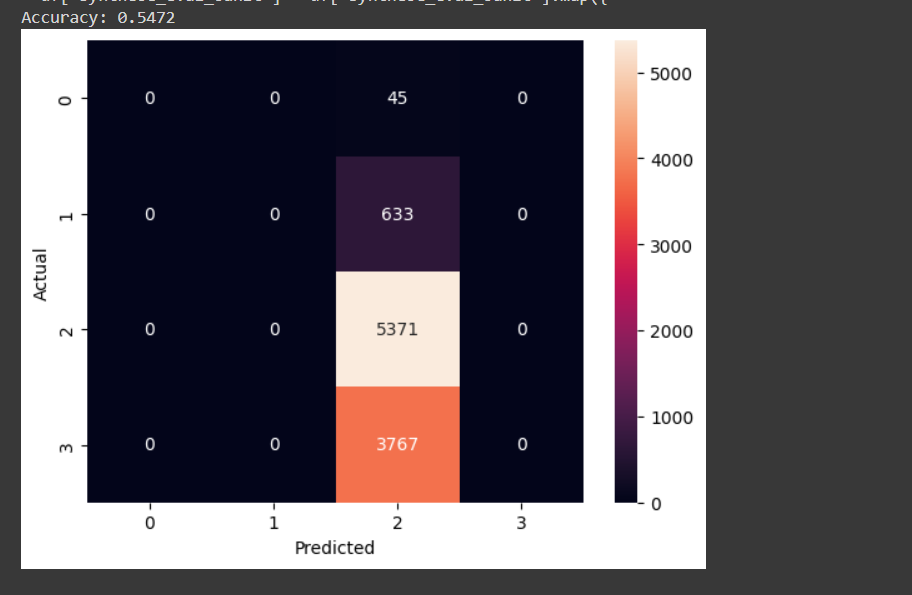
## Modèles

* Régression linéaire

Pour la régression linéaire, nous avons choisis de prédire la note de satisfaction selon le SIRET, le filtre et le type d’activité.



* Régression logistique



* Random forest

