**Les hyper-paramètres pour la partie du Machine learning**

Pour réaliser la partie machine learning, nous avons décidé d’utiliser trois approches :

* Le SVM (Support Vector Machine)
* RandomForest
* K-NN

1. SVM

Le modèle SVC a plusieurs hyperparamètres qui doivent être définis avant l'entraînement du modèle. Voici une liste des hyperparamètres les plus importants :

* C : C'est un hyperparamètre de régularisation qui contrôle la marge d'erreur. Une

valeur plus élevée de C permet au modèle de tolérer moins d'erreurs de classification sur les données d'entraînement. En d'autres termes, une valeur plus élevée de C conduit à une frontière de décision plus complexe et plus flexible, tandis qu'une valeur plus faible de C conduit à une frontière de décision plus simple et moins flexible.

* kernel : Le noyau utilisé pour la transformation des données dans un espace de

dimension supérieure. Les noyaux les plus couramment utilisés sont le noyau linéaire, le noyau polynomial et le noyau RBF.

* gamma : C'est un hyperparamètre de régularisation qui contrôle la flexibilité de la

frontière de décision. Une valeur plus élevée de gamma conduit à une frontière de décision plus complexe et plus flexible, tandis qu'une valeur plus faible de gamma conduit à une frontière de décision plus simple et moins flexible.

1. RandomForest

* n\_estimators : Le nombre d'arbres dans la forêt. Plus il y a d'arbres, meilleur est le

modèle.

* max\_depth : La profondeur maximale des arbres. Si non spécifié, les arbres sont développés jusqu'à ce que toutes les feuilles soient pures ou contiennent moins d'échantillons que le paramètre min\_samples\_split.

1. KNN

* n\_neighbors : Le nombre de voisins à utiliser. C'est le paramètre 'K' dans KNN qui détermine le nombre de voisins les plus proches à examiner pour la prédiction.
* weights : Fonction de poids utilisée dans les prédictions. Les options incluent 'uniform' (tous les points dans chaque voisinage sont pondérés également) et 'distance' (les voisins les plus proches d'un point de requête auront une influence plus grande).
* algorithm : L'algorithme utilisé pour calculer les voisins les plus proches. Les options incluent 'auto', 'ball\_tree', 'kd\_tree', et 'brute'.
* metric : La métrique de distance à utiliser pour l'arbre. Les options standards incluent 'minkowski', 'euclidean', 'manhattan', etc.