

# Analyses et résultats des données de l'égypte : Le fichier Egypte\_period\_prediction\_V2.csv

chahrazed labba

le 17 Mai 2023

Ce document décrit les données utilisées et présente les résultats des prédictions obtenues en utilisant les algorithmes de machine learning suivants : RF, svm et xGboost.

Données utilisées : une modification de la classe d'âge concernant 166 qui bascule dans les 5-12 ans (initialement sup 12 ans).

Il n'y a pas de changement dans les résultats lorsque l'on modifie les intervalles d'âge

## 1 Prédire la période temporelle avec 17 étiquettes

	A totale	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16
RF (%)	35.23	68.33	45.16	39.34	24.48	18.18	0.0	14.28	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SVM (%)	33.01	58.33	45.16	29.50	20.40	27.27	6.66	21.42	25.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Xgboost(%)	35.55	65.0	50.0	34.42	30.61	13.63	0.0	21.42	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 1: Accuracy globale et accuracy par class label

- df\_results\_SVM\_17\_labels\_LOO\_V2.csv (contient les prédictions faites par SVM)
- df\_results\_RF\_17\_labels\_LOO\_V2.csv (contient les prédictions faites par RF)
- df\_results\_xgboost\_17\_labels\_V2.csv (contient les prédictions faites par Xgboost)

## 2 Prédire la période temporelle avec 5 étiquettes

- df\_results\_SVM\_5\_labels\_LOO\_V2.csv (contient les prédictions faites par SVM)
- df\_results\_RF\_5\_labels\_LOO\_V2.csv (contient les prédictions faites par RF)
- df\_results\_xgboost\_5\_labels\_V2.csv (contient les prédictions faites par Xgboost)

	Accuracy	0	1	2	3	4	5
RF	78.41	99.55	35.0	0.0	0.0	0.0	0.0
SVM	77.46	92.95	45.0	27.27	0.0	0.0	0.0
Xgboost	78.09	93.83	46.66	22.72	0.0	0.0	0.0

Table 2: Accuracy globale et accuracy par class label

### 3 Résultat de clustering en utilisant k-means

Le clustering donne lieu à deux clusters.

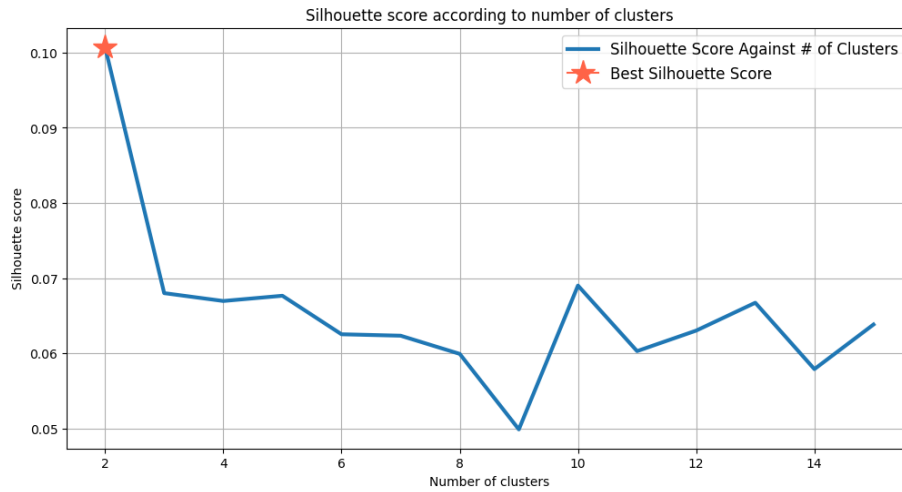


Figure 1: Nombre de clusters

Comme K-means est une boîte noire et que nous ne pouvons pas savoir quelles caractéristiques ont été utilisées pour réaliser le clustering, j'ai appliqué un algorithme de classification (Random Forest), puis la lib Shap pour expliquer le clustering.

Le fichier Egypt\_Cluster\_Data\_315\_V2.csv contient les résultats de clustering.

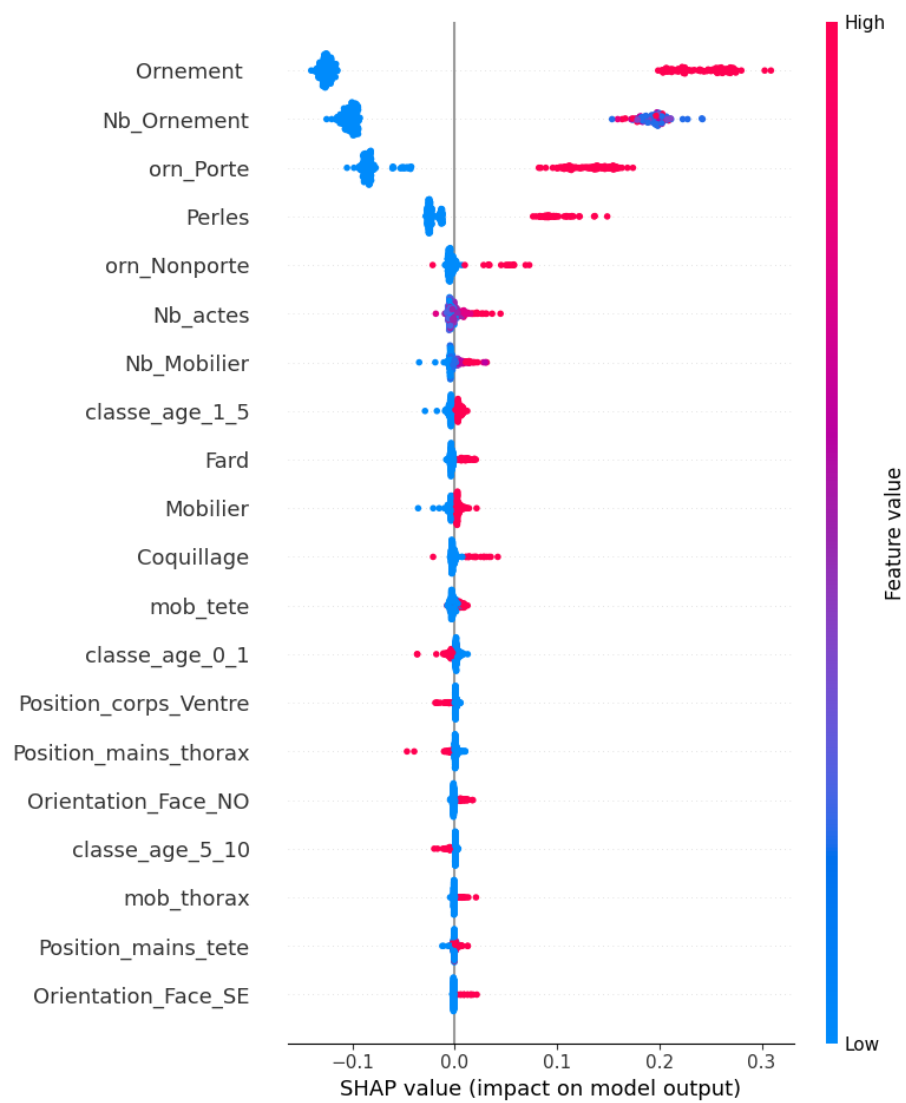


Figure 2: Cluster 0

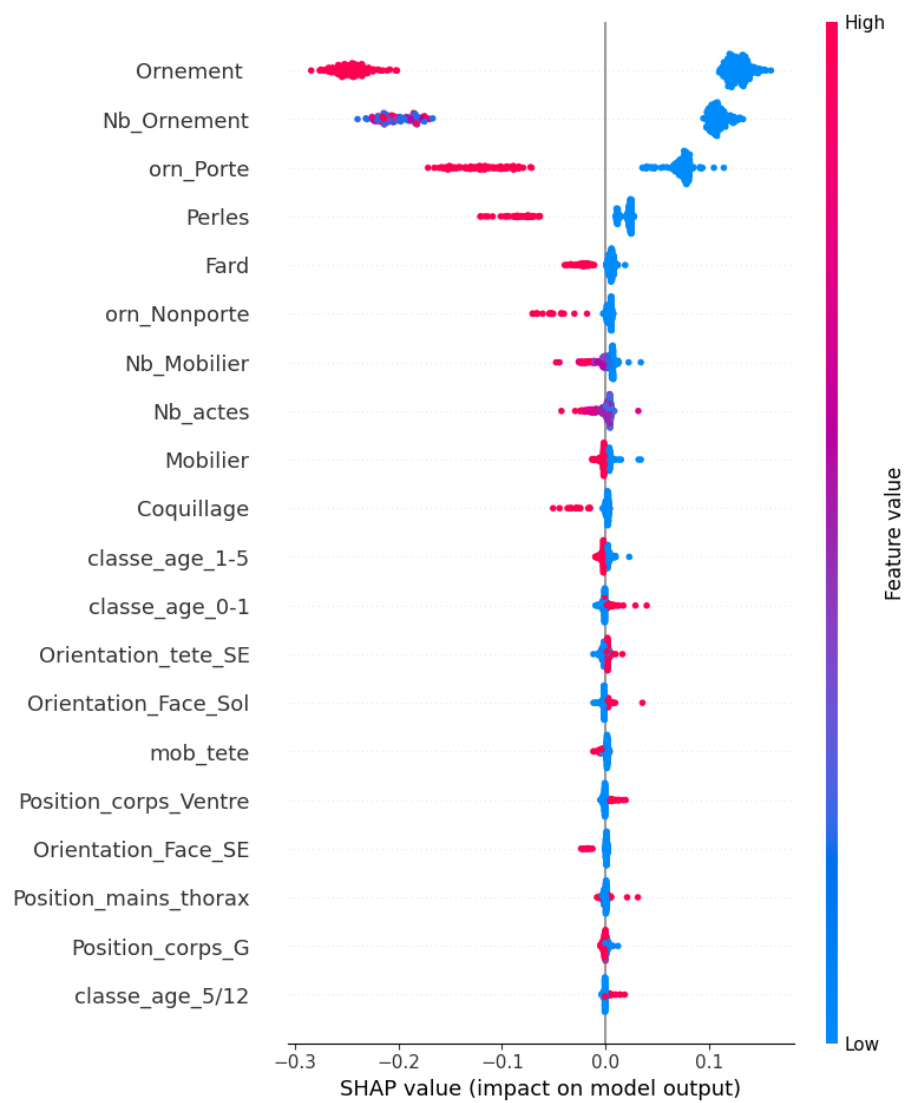


Figure 3: Cluster 1