

Analyses et résultats des données de l'égypte : Prédire le sexe

chahrazed labba

le 17 Mai 2023

Ce document résume les résultats relatifs à la prédiction du sexe en utilisant des données d'entraînement de 27 sujets, dont 8 hommes et 19 femmes.

J'ai utilisé trois algorithmes d'apprentissage automatique : RF, SVM et XGboost. Les algorithmes ont été testés en utilisant la validation croisée et en particulier la méthode "leave one out" pour vérifier leurs performances.

Dans un deuxième temps, chaque algorithme est entraîné sur l'ensemble des données d'entraînement (en une seule fois), puis le modèle obtenu est utilisé pour prédire le sexe de l'individu.

Pour chaque algorithme, j'ai utilisé le Shap lib pour expliquer quelles caractéristiques ont été utilisées pour décider du sexe.

Les étiquettes de classes : 0 pour Homme, et 1 pour Femme

1 Prédire le Sexe avec RF

1.1 Leave one Out

- Accuracy Globale : 66.66 %
- Accuracy pour class Label 0 (homme) : 12.5 %
- Accuracy pour class Label 1 (femme) : 89.47 %

1.2 Prédiction de Indet

Le fichier prediction_sex_RF.csv' contient les résultats de prédiction du sexe en utilisant RF.

Les deux figures (1-a, 1-b) suivantes présentent les caractéristiques utilisées par RF pour faire les prédictions.

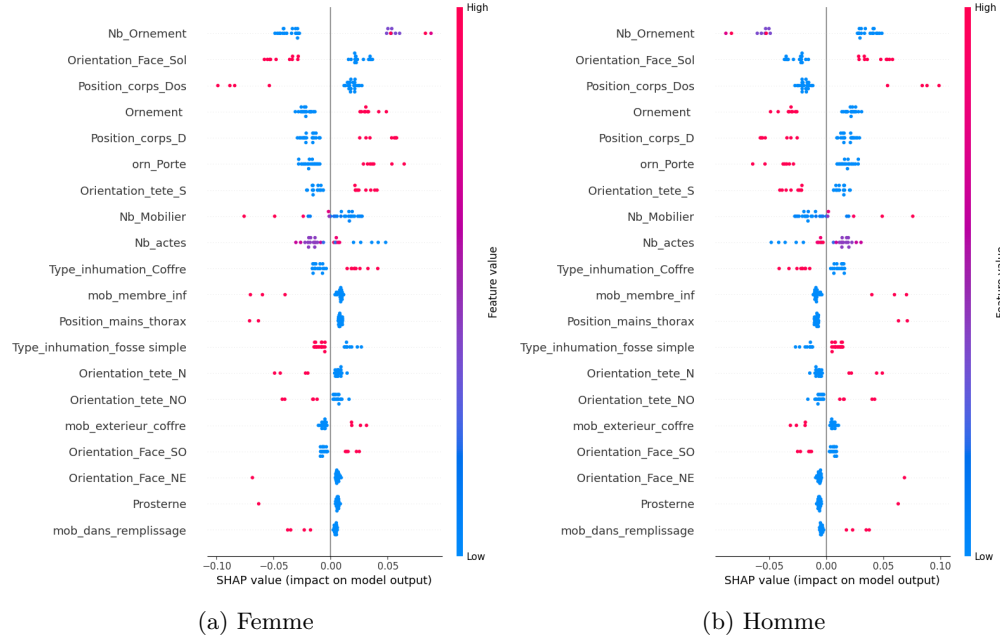


Figure 1: Caractéristiques utilisées pour la prédiction du sexe par RF

2 Prédire le Sexe avec SVM

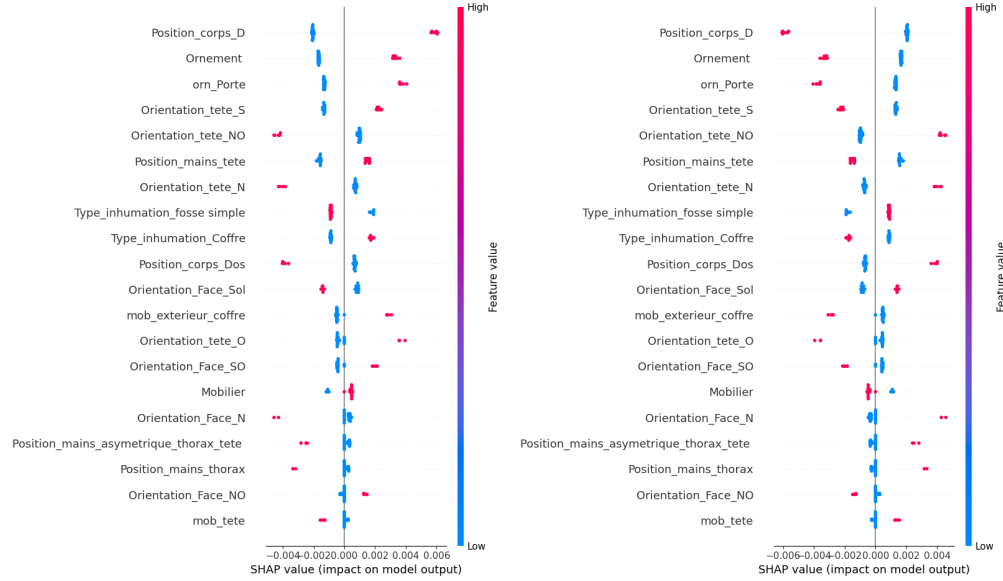
2.1 Leave one Out

- Accuracy Globale : 55.55 %
- Accuracy pour class Label 0 (homme) : 12.5 %
- Accuracy pour class Label 1 (femme) : 73.68 %

2.2 Prédiction de Indet

Le fichier 'prediction_sex.SVM.csv' contient les résultats de prédiction du sexe en utilisant SVM.

Les deux figures suivantes présentent les caractéristiques utilisées par SVM pour faire les prédictions.



(a) Femme

(b) Homme

Figure 2: Caractéristiques utilisées pour la prédiction du sexe par SVM

3 Prédire le Sexe avec XGboost

3.1 Leave one Out

- Accuracy Globale : 66.66 %
- Accuracy pour class Label 0 (homme) : 25.0 %
- Accuracy pour class Label 1 (femme) : 84.21 %

3.2 Prédiction de Indet

Le fichier 'prediction_sex_XGboost.csv' contient les résultats de prédiction du sexe en utilisant XGboost.

Les deux figures suivantes présentent les caractéristiques utilisées par XGboost pour faire les prédictions.

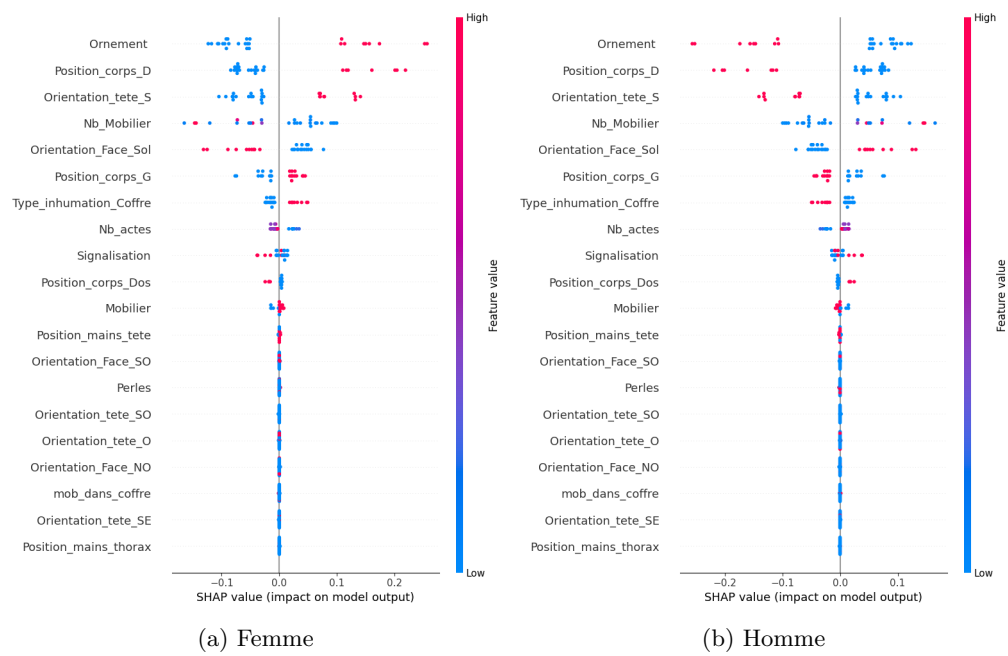


Figure 3: Caractéristiques utilisées pour la prédiction du sexe par XGboost