

# Rapport d'évaluation du projet 2 4algl4a -- 2024-2025

## -- NVS

**60283 Salim Ben Ali**

remis

### défense

ok

### rapport

#### auteur

- ko : absent

#### titre

ok

#### sources Python

ok

#### sorties des sources

##### textuelles

ok, sauf pour iris où les données produites dans le rapport sont partielles

##### graphique

ok, sauf :

- pour iris où une figure est dupliquée
- pour mall où un graphe inertie vs. k est produit par le code mais pas inclus au rapport

#### analyse / nombre idéal de regroupements

ok :

- les valeurs optimales de k sur base de la silhouette (et de l'ari si possible) sont fournies

ko :

- pour iris : la valeur optimale annoncée est  $k = 3$  : ok pour ari, mais silhouette donne 2 et l'étudiant n'en parle pas

#### 4 jeux de données

ok

#### découpe en semaines

non

#### autre

(void)

## **sources**

### **pré-traitements**

#### **valeurs manquantes**

(void)

#### **valeurs aberrantes**

(void)

#### **données mall injection valeurs manquantes**

ko :

- pas fait

#### **imputation**

(void)

#### **réduction de dimension**

ok :

- mall : retrait d'id

#### **mise à échelle**

ok :

- StandardScaler() : iris, wine, breast, mall

#### **algorithme kmeans**

##### **recherche du meilleur k entre 2 et 5 avec erreur carrée, silhouette et davies-bouldin**

ok :

- métriques calculées : inertie, silhouette, davies-bouldin, ari (sauf pour mall)

ko :

- pas d'algorithme kmeans pour breast
- pas de détection automatique du meilleur k

#### **algorithme kmeans++**

##### **recherche du meilleur k entre 2 et 5 avec erreur carrée, silhouette et davies-bouldin**

ok :

- métriques calculées : inertie, silhouette, davies-bouldin, ari (sauf pour mall)

ko :

- pas de détection automatique du meilleur k

#### **autre**

- pour breast : pas de kmeans mais knn (k nearest neighbours)
- pour breast : séparation en données d'entraînement et de tests
- pour mall : remplacement des données textuelles (sexe : Male et Female) en numérique (0 et 1)
- pour mall : comparaison des métriques avec et sans mise à l'échelle

- sauvegarde des figures en png

**conclusion**

- rapport : comme demandé à de petites exceptions près
  - code : ok
-