

# Projet de classification non supervisée : détection d'activités humaines

à rendre au plus tard le XX/11/2025 (FI) et le XX+7/11/2025 (FA)

## 1 Objectifs

L'objectif de ce projet est d'appliquer des algorithmes de classification non supervisée afin d'identifier, à partir de données collectées par des smartphones, des classes correspondant aux activités physiques d'une personne : (1) marche, (2) montée d'escaliers, (3) descente d'escaliers, (4) position assise, (5) position debout, (6) position allongée.

## 2 Acquisition des données

Pour effectuer la collecte des données, une personne a suivi un parcours contenant les différentes activités listées ci-dessus. Lors du parcours, 9 variables ont été enregistrées toutes les 0.02 seconde : 3 accélérations réelles (suivant les axes x, y, z), 3 accélérations estimées (suivant les axes x, y, z) et 3 vitesses (suivant les axes x, y, z). Les données utilisées sont issues des travaux [1]. Un extrait video des expériences réalisées peut être visualisé via le lien [http://www.youtube.com/watch?v=XOEN9W05\\_4A](http://www.youtube.com/watch?v=XOEN9W05_4A).

## 3 Description des données

Pour faciliter l'analyse, les données ont été découpées en fenêtres temporelles de 128 observations (chaque fenêtre correspond à un extrait de 2.5 secondes). Ce choix méthodologique repose sur l'idée qu'il est plus pertinent d'identifier les activités sur des intervalles de temps plutôt que sur des instants ponctuels.

Ces données sont rangées dans **9 tableaux de taille identique** ( $347 \times 128$ ) dont chaque ligne correspond à une même fenêtre temporelle de 128 observations. On dispose donc de 347 fenêtres successives.

accm_x	accélération mesurée suivant l'axe x (appelée accélération totale)
accm_y	accélération mesurée suivant l'axe y
accm_z	accélération mesurée suivant l'axe z
acce_x	accélération estimée suivant l'axe x (avec suppression des effets liés à la gravité)
acce_y	accélération estimée suivant l'axe y
acce_z	accélération estimée suivant l'axe z
vit_x	vitesse suivant l'axe x
vit_y	vitesse suivant l'axe y
vit_z	vitesse suivant l'axe z

On dispose également des vraies classes (**lab**) qui pourront servir à évaluer les méthodes proposées. Au total, il y a **10 fichiers de données** à exploiter dans ce projet. La figure 1 montre la mesure d'accélération sur deux fenêtres de temps : la fenêtre 2 et la fenêtre 100 (on rappelle que chaque fenêtre comporte 128 mesures).

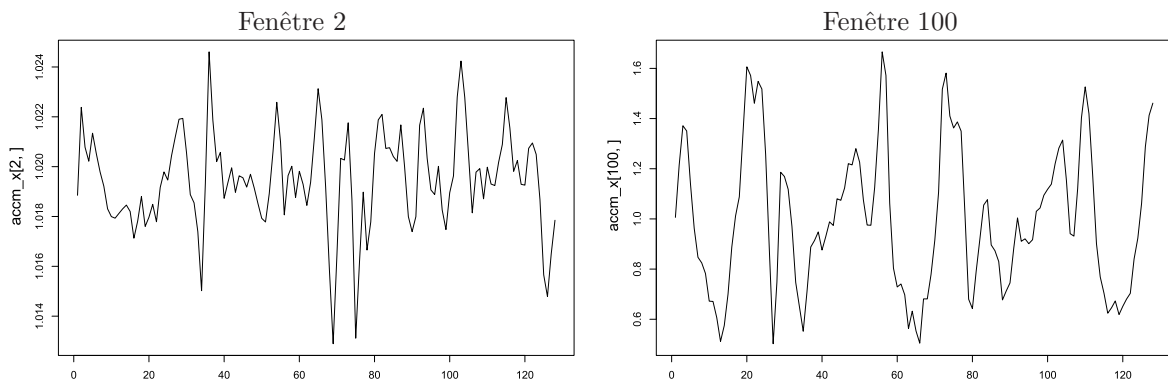


FIGURE 1 – Accélération longitudinale (accm x) pour deux fenêtres de temps

La figure 2 montre l'ensemble des 9 variables (l'axe des abscisses indice les 347 fenêtres de temps, l'axe des ordonnées indice les 128 mesures par fenêtre de temps, et la couleur correspond à la valeur numérique de la variable considérée).

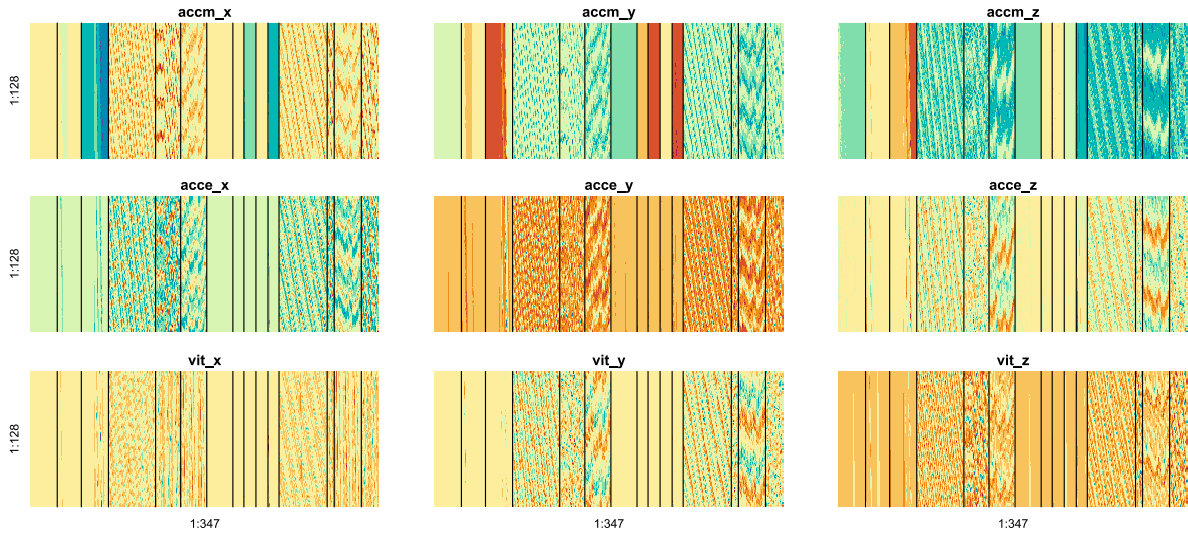


FIGURE 2 – Représentation de l'ensemble des données, variable par variable : l'axe des abscisses indice les 347 fenêtres de temps, l'axe des ordonnées indice les 128 mesures par fenêtre de temps, et la couleur correspond à la valeur numérique de la variable considérée, du bleu (faible), au rouge (élevé) ; les lignes verticales délimitent les classes

## 4 Travail à effectuer

Réaliser une classification automatique des 347 individus (fenêtres temporelles) pour détecter automatiquement les différentes classes d'activité. Pour cela, il est demandé d'appliquer et de comparer au moins trois algorithmes ou méthodes étudiés en cours. *Vous pourrez utiliser des distances adaptées aux séries temporelles, et, si cela vous semble pertinent, procéder à une transformation préalable des données, par exemple en remplaçant chaque fenêtre par un ensemble de caractéristiques telles que la moyenne, la variance, les moments d'ordre  $p$ , le minimum, le maximum, coefficients d'auto-régression, de Fourier, ...*

Après avoir réalisé la classification, analyser les résultats obtenus.

## 5 Documents à rendre

Deux fichiers : un rapport en pdf n'excédant pas 15 pages et un fichier contenant le code R ou Python associé.

## Référence

- [1] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz (2013). A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges.