

Comparaison de méthodes de classification multivues à l'aide du logiciel R : application à des données archéologiques

CV à adresser à : lise.bellanger@univ-ubs.fr

Descriptif de l'offre

Depuis de nombreuses années, le Laboratoire Archéologie et Territoires (CITERES-LAT¹, UMR 7324, CNRS – Université de Tours) collabore avec des équipes de statistique afin de concevoir des outils statistiques destinés à l'étude de la chronologie. Une méthode de Classification Ascendante Hiérarchique (CAH) par compromis a notamment été proposée afin de déterminer une partition consensuelle entre deux sources d'information (ex. céramique et stratigraphie) (Bellanger, Coulon, Husi 2021). Cette méthode est désormais implémentée dans le package R [SPARTAAS](#) (*Statistical Pattern Recognition and daTing using Archeological Artefacts assemblages*) (Coulon, Bellanger, Husi 2023).

Le stage volontaire de niveau M1 proposé s'inscrit dans le cadre du projet [Arctools](#) (OPEN CNRS) dont l'objectif est d'enrichir le package R SPARTAAS en finalisant sa brique chronologique. Le stage sera encadré en collaboration avec le laboratoire Lab-STICC (UMR 6285, Université Bretagne Sud) pour l'aspect méthodologique en statistique.

L'objectif du stage est de mettre en œuvre sous R différentes méthodes de classification non supervisée multivues, puis de les évaluer et de les comparer à l'approche CAH par compromis développée. Les méthodes seront appliquées à des jeux de données simulés ainsi qu'à des données archéologiques réelles, en particulier un corpus céramique d'Angkor Thom (Cambodge, 9^e–15^e s.), déjà organisé au sein d'une base de données.

Profil recherché

- Master 1 Pro/Recherche ou Ecole d'ingénieur en statistique
- Connaissances indispensables des notions de base en classification non supervisée
- Maîtrise du langage R ainsi que des outils R Markdown, Shiny, Quarto et Latex pour la création de documents, interfaces et rapports
- Intérêt pour les applications de la statistique
- Autonomie, rigueur, capacités de vulgarisation fortement recommandées
- Intérêt pour le travail interdisciplinaire

Précisions pratiques

Lieu	Laboratoire Archéologie et Territoires (UMR CITERES-LAT), Tours
Durée	2 mois (date de début à discuter)
Encadrement professionnel	Philippe Husi (UMR 7324 CITERES-LAT ; philippe.husi@univ-tours.fr) et Arthur Coulon (UMR 7324 CITERES-LAT ; arthur.coulon@univ-tours.fr) Lise Bellanger (UMR 6285 Laboratoire Lab-STICC, lise.bellanger@univ-ubs.fr)

Références

Bellanger, L., Coulon, A. and Husi, P. (2021). PerioClust: A simple hierarchical agglomerative lustering approach including constraints (p. 1–8). Springer International Publishing.

https://doi.org/10.1007/978-3-030-60104-1_1

Coulon A., Bellanger L., Husi P. 2023. – SPARTAAS: Statistical Pattern Recognition and daTing using Archaeological Artefacts assemblageS. R package version 1.2.1, <https://CRAN.R-project.org/package=SPARTAAS>.

Hulot, A., J. Chiquet, F. Jaffrézic et al. (2020). Fast tree aggregation for consensus hierarchical clustering. BMC Bioinformatics 21, 120. <https://doi.org/10.1186/s12859-020-3453-6>

¹ Voir <http://citeres.univ-tours.fr/>