

Classification Ascendante Hiérarchique par compromis : extension, étude de robustesse et application à des données archéologiques

CV, lettre de motivation et coordonnées d'au moins une référence 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 de niveau M2 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.

Les missions confiées à la personne recrutée porteront sur l'extension de la CAH par compromis à plus de deux sources d'informations et incluront :

- (i) la validation d'une méthode d'optimisation du critère permettant l'obtention du compromis, ainsi que l'étude des propriétés et de la robustesse de cette approche ;
- (ii) la comparaison avec différentes méthodes de classification non supervisée multivues ;
- (iii) l'intégration de cette extension dans une fonction R dédiée du package SPARTAAS.

Différents corpus de données archéologiques, provenant de l'étude de la céramique - source matérielle omniprésente en archéologie - pourront être mobilisés (fouilles archéologiques réalisées à Angkor Thom, capitale de l'empire khmer (9^e et le 15^e s.) ou dans le bassin de la Loire Moyenne).

Profil recherché

- Master 2 Pro/Recherche ou Ecole d'ingénieurs 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 ; Intérêt pour le travail interdisciplinaire
- Autonomie, rigueur, capacités de vulgarisation fortement recommandées

Précisions pratiques

Lieu	Laboratoire Archéologie et Territoires (UMR CITERES-LAT), Tours
Durée	6 mois (date de début à discuter)
Gratification	Environ 600€/mois
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/>