

Créer son propre package d’extension `{recipes}`: retour d’expérience de `{scimo}`

Antoine Bichat*

Julie Aubert†

Résumé

L’avènement des technologies de séquençage à haut débit a entraîné une augmentation massive de la production de données omiques : génomiques, transcriptomiques, protéomiques, métagénomiques, et bien plus encore.

Pour explorer et analyser efficacement les données omiques, des étapes de prétraitement adaptées sont requises, comme la normalisation, la sélection et l’agrégation de variables (Perez-Riverol et al. (2017)). Cependant, ces méthodes spécifiques ne sont pas disponibles nativement dans le package `{recipes}` (Kuhn, Wickham, and Hvitfeldt (2024)). Nous avons alors développé un package d’extension, `{scimo}`, conçu pour intégrer ces étapes dans l’écosystème `{tidymodels}` (Kuhn and Wickham (2020)).

`{scimo}` propose une série d’étapes de prétraitement adaptées à l’analyse des données omiques, tout en restant adaptable à d’autres types de données. Il est disponible à cette adresse : <https://github.com/abichat/scimo>

Mots-clés : Tidymodels - Package - Statistique - Biostatistique - Données omiques

Développement

Au cours de cette présentation, nous introduirons les fonctions principales du package `{scimo}` et expliquerons comment créer son propre package d’extension `{recipes}`. Nous insisterons sur les difficultés rencontrées au cours du développement.

Références

- Kuhn, Max, and Hadley Wickham. 2020. *Tidymodels: A Collection of Packages for Modeling and Machine Learning Using Tidyverse Principles*. <https://www.tidymodels.org>.
- Kuhn, Max, Hadley Wickham, and Emil Hvitfeldt. 2024. *Recipes: Preprocessing and Feature Engineering Steps for Modeling*. <https://CRAN.R-project.org/package=recipes>.
- Perez-Riverol, Yasset, Max Kuhn, Juan Antonio Vizcaino, Marc-Phillip Hitz, and Enrique Audain. 2017. “Accurate and Fast Feature Selection Workflow for High-Dimensional Omics Data.” *PloS One* 12 (12): e0189875.

*Les Laboratoires Servier, antoine.bichat@servier.com

†MIA Paris-Saclay, julie.aubert@agroparistech.fr