

M2 Research internship Machine learning and survival in Julia.

Context

Machine learning and deep learning enrich survival analysis by managing complex relationships between variables ¹, but also by proposing models that can handle multimodal data, such as imaging and biological data ². Most of these models are implemented in Python or R (e.g., ^{3,4}). R, which is less efficient with large data sets, is also less suitable than Python for the latest deep learning tools. Conversely, the Python community for survival analysis is growing rapidly, but remains less mature and developed than the R community.

Objectives

The main objective of the internship is to develop a new library in the -- rising -- programming language Julia to implement machine and deep learning survival algorithms efficiently. The JuliaSurv organization⁵ demonstrate Julia's advancement in survival analysis, while libraries such as MLJ.jl, Flux.jl and Lux.jl highlight its capabilities in deep learning. Julia offers notable benefits in terms of performance, integration with deep learning and ease of development.

The communication and publication of our results will be an integral part of the work. This is also a very good opportunity to learn Julia, a new programming language that is gaining traction in both academic and industry.

Candidate's profile

The following are strictly required:

- A master 2 level in statistics, computer sciences, data science or related fields
- Very good programming skills
- Knowledge of git and ability to read R and Python
- Very good written English

The following are not required but will be nice bonuses:

Knowledge of standard survival analysis, of standard machine and deep learning and of the Julia language.

Additional information

- **Length:** 4 to 6 months.
- **Location:** The internship will take place at the SESSTIM on the Faculté des Sciences Médicales et Paramédicales in Marseille, France.
- **Wages:** Regulatory internship stipend at the Université Aix-Marseille (4€35/hour for 35h/week, about 600€/month)
- **Advisors:** Oskar Laverny (MCF in statistics, SESSTIM, AMU) and Nathalie Grafféo (IR in biostatistics, SESSTIM, INSERM)
- **Contact details:** Please send your application, containing a resume and a cover letter, to oskar.laverny@univ-amu.fr and nathalie.graffeo@univ-amu.fr with [StageMLSurv2025] in the mail's object.

References

1. Katzman, J. L. *et al.* DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med. Res. Methodol.* **18**, 24 (2018).
2. Tripathi, A., Waqas, A., Yilmaz, Y. & Rasool, G. Abstract 4905: Multimodal transformer model improves survival prediction in lung cancer compared to unimodal approaches. *Cancer Res.* **84**, 4905 (2024).
3. Pölsterl, S. scikit-survival: A Library for Time-to-Event Analysis Built on Top of scikit-learn. *J. Mach. Learn. Res.* **21**, 1–6 (2020).
4. Sonabend, R., Király, F. J., Bender, A., Bischl, B. & Lang, M. mlr3proba: an R package for machine learning in survival analysis. *Bioinformatics* **37**, 2789–2791 (2021).
5. Alhajal, R. & Laverny, O. NetSurvival.jl: A glimpse into relative survival analysis with Julia. Preprint at <https://doi.org/10.48550/arXiv.2408.15655> (2024).