

## **TITLE: Personalized Health Care**

### **Description**

This lab will give an overview about survival analysis in observational studies. The goal is to develop a prediction model based on high-dimensional patient profiles. Machine Learning methods will be used to challenge more traditional regression models or clinical risk stratification.

### **Duration**

20 hours (5 slots of 4 hours)

### **Calendar** (tentative) 1430 – 1830:

- May 4
- May 11
- May 18
- May 25
- June 1

**Eligible Students:** max. 15 students from DSE 1<sup>st</sup> and 2<sup>nd</sup> year.

### **Main Content**

- Develop a predictive model according to your preferred ML algorithm (Lasso, Random Forest, Neural Network, etc.) based on clinical, molecular and gene expression data. The response of the model is a time to event outcome, therefore a regression model able to deal with a censored outcome must be considered.
- Evaluate the predictive performance of the model. The evaluation phase is extremely important and have to be conducted accounting for the specific characteristics of survival data.

The R package `mlr3` has a section dedicated to prediction models with survival data and can be used as a platform for developing the project.

Students will work in groups. Each group will do a final presentation of the developed approach and of the results obtained.

### References:

- <https://mlr3book.mlr-org.com>
- Ishwaran H, Kogalur UB, Gorodeski EZ, Minn AJ and Lauer MS (2010). "High-dimensional variable selection for survival data." *Journal of the American Statistical Association*, 105(489), 205-217.
- Binder, H. and Schumacher, M. (2008). Allowing for mandatory covariates in boosting estimation of sparse high-dimensional survival models. *BMC Bioinformatics*. 9:14.
- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. *An Introduction to Statistical Learning: with Applications in R*. New York :Springer, 2013.
- Wang, J., Heng, Y.J., Eliassen, A.H. et al. Alcohol consumption and breast tumor gene expression. *Breast Cancer Res* 19, 108 (2017).

### Reference for survival analysis:

Theory: BULL, K. and SPIEGELHALTER, D.J. (1997), Tutorial in biostatistics: survival analysis in observational studies. *Statist. Med.*, 16: 1041-1074.

### **How to apply**

Send an email to [dse@unimi.it](mailto:dse@unimi.it) with your CV, your DSE career self-certification and cover letter explaining why you want to join this lab and how you meet the requirements.

### **Deadline**

3rd March 2021 at 23:59 (CET).