scikit-fda: Functional Data Analysis in Python

Carlos Ramos Carreño

(Universidad Autónoma de Madrid)

14.04.2021, 16.00 (c.t.)

Online via Zoom
(Meeting-ID: 913-2473-4411; Password: StatsCol21)

In this talk I will introduce scikit-fda, a Python library for Functional Data Analysis. This library integrates itself in the SciPy ecosystem of Python scientific packages, and in particular can be used alongside the Machine Learning library scikit-learn, reusing some of the tools available for it. Along with flexible representations of functional data observations, including curves, surfaces or vector-valued functions, scikit-fda provides tools to preprocess functional data (smoothing, registration, dimensionality reduction), aid in exploratory analysis (providing outlier detection, visualization or depth computation tools), perform statistical analysis such as functional ANOVA, or apply machine learning methods (classification, regression, clustering). I will give an overview of these functionalities, showing their usage and how they can be combined with existing tools, such as the hyperparameter optimization tools already available in scikit-learn, in order to perform a complete analysis of functional data.

Biography:
Carlos Ramos Carreño is a Computer Science PhD student from the GAA (Grupo de Aprendizaje Automático, translated as Machine Learning Group) in Universidad Autónoma de Madrid. He has Bachelor and Master degrees in both Mathematics and Computer Science Engineering. His area of research is Functional Data Analysis applied to Machine Learning. As part of his thesis he is currently building and maintaining scikit-fda, which is intended to be the go-to library for Functional Data Analysis in Python. He has also contributed several open source tools to the Python scientific ecosystem, such as dcor, a package for computing distance correlation and related E-statistics, and rdata, a package to import data in R to Python. He is also active in the Python scientific community and has sucessfully contributed to popular scientific packages such as Scipy or Sacred.