

Statistical data analysis and Machine learning

The aim of this lecture is twofold, provide the essential knowledge for efficient statistical data analysis and give an introduction to modern machine learning methods. These techniques are widely used in many scientific areas and their importance is continuously growing. Hands-on sessions come as a support of the lectures with the objective that skills acquired can be easily and efficiently used afterwards. This course is addressed to all PhD students of the EDSF, no specific knowledge of statistical analysis, machine learning or Python language is required.

Lecturers: Emmanuel BUSATO and Julien DONINI

Part I – Statistical data analysis

- **Fundamental concepts:** probabilities, statistical model, likelihood function
- **Statistical inference:** parameter estimation, hypothesis tests, frequentist and bayesian approaches
- **Hands-on session** with Python: data analysis, fitting a statistical model

Part II – Introduction to machine learning and neural networks

- **Basic concepts:** linear regression and classification, supervised and unsupervised learning
- **Introduction to neural networks:** constructing and training a network, review of popular neural networks used today
- **Hands-on session:** setting up a neural network using basic tools (python, numpy) and advanced machine learning libraries (pytorch)