Università degli Studi della Campania *Luigi Vanvitelli*
*Dipartimento di Matematica e Fisica*

Istituto Nazionale di alta Matematica "Francesco Severi"

Università degli Studi della Campania "Luigi Vanvitelli"
Dipartimento di Matematica e Fisica

***PhD Program in Mathematics, Physics and Applications to Engineering***

# COURSE "DATA SCIENCE: AN ALGORITHMIC APPROACH"

**Prof. Erkki Somersalo**
**Case Western Reserve University, Cleveland, Ohio, USA**

## Course description

In order to extract pertinent information out of the enormous amount of data that are currently available, it is necessary to use sophisticated computational tools that are developed to reduce, visualize or otherwise compress the data in a form that is understandable to humans. This course focuses on a number of fundamental principles and algorithms that are widely used in modern data science, from analyzing scientific data to feature extraction from text and image data.

Rather than discussing ready-made black box algorithm packages, the emphasis of this course is on understanding the basic ideas behind the algorithms, as well as on the implementation of them from those first principles. The topics include data reduction and visualization, clustering, classification algorithms, text document analysis, texture analysis of images. Methodologically, the course is built mostly on basic linear algebraic methods and does not require a background in statistics. It provides the tools for the students to build the algorithms from scratch. The algorithms discussed in this course include principal component analysis (PCA), linear discriminant analysis (LDA), k-means and k-medoids, tree-based classification methods, support vector machine (SVM), self-organizing maps (SOM), query matching and page ranking. Extensive numerical examples using real world data sets are discussed during the course. The programming language is Matlab.

***Erkki Somersalo*** *is a professor in the Department of Mathematics, Applied Mathematics and Statistics at Case Western Reserve University, with a background in analysis and partial differential equations. His research interests include computational inverse problems, with an emphasis on Bayesian methods, and applications to a wide range of areas, particularly to medical imaging. He is a founding member of the Finnish Inverse Problems Society and a member of the Finnish Academy of Sciences and Letters and the Society for Industrial and Applied Mathematics (SIAM).*

The course (16 hours) starts on May 4, 2021, and ends on June 3. Interested PhD students must send an email to **labinf.dmf@unicampania.it** to be admitted. The course lectures will be also posted on the **INdAM YouTube channel**.