



Giuliano Armano is associate professor of Information Processing Systems at the Faculty of Engineering, University of Cagliari. Member of the Dept. of Electrical and Electronic Engineering since 1993, he is team leader of the “Intelligent Agents and Soft-Computing” group (<http://www.iascgroup.it>). With a background on expert systems and machine learning, over the years he has investigated various kinds of learning strategies and techniques. His current research interests are focused on ensemble strategies, with particular emphasis on *mixtures of experts* and *hierarchical classification*. Additional research interests of prof. Armano are on performance measures for classifier systems, complex networks, and GPUs. These topics have been extensively experimented on

various application fields, in particular *bioinformatics*, *information retrieval* and *text categorization*. He is currently teaching *Elements of Bioinformatics*, for the degree program in Biomedical Engineering, and *Object-Oriented Programming Languages*, for the degree program in Electrical and Electronic Engineering. Since 2003, he also teaches a series of yearly seminars on the *Python language*. Starting from 1999, prof. Armano has been scientific director of several research projects, for a total of over one million euros. These projects include funding from the local government of Sardinia, from the Italian Ministry of Research, and private funding as well. Author of over 150 scientific publications (among which about 50 international journals), prof. Armano has also been guest editor of some journal special issues, including “Multi-Agent Systems for Medicine, Computational Biology, and Bioinformatics” (MultiAgent and Grid Systems Journal) and “GRID, Web Services, Software Agents, and Ontology Applications for Life Sciences” (IEEE Trans. on Nanobiosciences). Prof. Armano has also organized several workshops and conferences, including ISCC 2006, NETTAB 2006, AIIA 2008, KDWEB 2015 and KDWEB 2016.

Tutorial at BITS 2017 - Workflow management of bioinformatics data using Orange

Developed in Python, Orange is an open-source toolkit that provides great support for data visualization, machine learning and data mining tasks in terms of workflow management and embedded functionalities. With wellknown Python libraries that work behind the scenes (e.g., numpy, scipy and scikit-learn), Orange provides a nice and easy-to-use visual programming front-end aimed at performing interactive data visualization and analysis. In this tutorial, some examples of relevant problems that bioinformaticians may

have to solve during their everyday work are analyzed and handled using Orange –with specific focus on its ability of supporting the user in the definition of suitable workflows able to provide schematic and practical solutions.

Summary of the Talk

The tutorial will be focused on the following aspects:

- Introduction to data mining, data analysis and workflow management;
- Notes on the Orange application and its underlying technology;
- A look at the default tabs of Orange (i.e., “Data”, “Visualize”, “Classify”, “Regression”, and so on);
- Contents of the “Bioinformatics” tab (optional);

- Examples of relevant bioinformatics problems and their handling with Orange in terms of workflow management;
- How to embed user-defined libraries and algorithms in Orange.