UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE STATISTICHE

PHD COURSE IN STATISTICS

Specialist Course | Cycle XXXV

September, 2020 | Campus S. Caterina

Regression Modelling with Large Data Sets

Ioannis Kosmidis

University of Warwick and Alan Turing Institute

Thursday	**September 17**	**10.00 – 13.00**	**Room Benvenuti**
		14.00 – 16.00	**Room Benvenuti**
Friday	**September 18**	**10.00 – 13.00**	**Room Benvenuti**
		14.00 – 16.00	**Room Benvenuti**

Abstract

www.stat.unipd.it/fare-ricerca/courses-201920-xxxv-cycle

Regression Modelling with Large Data Sets

Ioannis Kosmidis
University of Warwick and Alan Turing Institute

Abstract

This short course aims to introduce core theory, methods and software tools for tackling regression problems that involve large data sets in terms of number of observations or in terms of explanatory variables. A key learning objective is to identify least squares as a core optimisation problem for fitting linear and generalised linear models, and classify the various methods for its solution or regularised versions of it in terms of complexity, memory usage and accuracy. Upon completion of the course, the participants will be able to describe a range of incremental bounded-memory algorithms and stochastic gradient descent algorithms for large regression problems, contrast them in terms of their relative merits, and use available software and tools to apply them on real data sets.