

Mathematics for Data Science

Claudio Agostinelli
claudio.agostinelli@unitn.it

Department of Mathematics
University of Trento

07 April 2017

Data Science

- **Data science** is an interdisciplinary field about processes and systems to extract knowledge or insights from data in various forms, either structured or unstructured, which is a continuation of some of the data analysis fields such as statistics, data mining, machine learning and predictive analytics. Wikipedia.
- The Sexiest Job of the 21st Century. **Data scientists** are the key to realizing the opportunities presented by big data. They bring structure to it, find compelling patterns in it, and advise executives on the implications for products, processes, and decisions. They find the story buried in the data and communicate it. And they don't just deliver reports: They get at the questions at the heart of problems and devise creative approaches to them ...” Harvard Business Review, October 2012
- McKinsey & Company projecting a global excess demand of 1.5 million new data scientists.

Mathematics for Data Science

- New York Times (2009): For Today's Graduate, Just One Word: Statistics; Hal Varian [Google's Chief Economist] (2009): "I keep saying that the sexy job in the next 10 years will be statisticians. And I'm not kidding."
- Wired (2009), "Hal Varian [Google's Chief Economist] believes that a new era is dawning for what you might call the dataratiand it's all about harnessing supply and demand. Whats ubiquitous and cheap? Varian asks. 'Data.' And what is scarce? The analytic ability to utilize that data. As a result, he believes that the kind of technical person who once would have wound up working for a hedge fund on Wall Street will now work at a firm whose business hinges on making smart, daring choices—decisions based on surprising results gleaned from algorithmic spelunking and executed **with the confidence that comes from really doing the math.**"

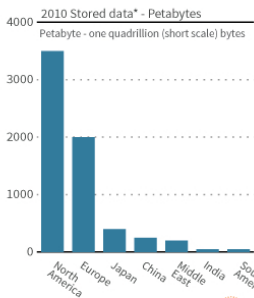
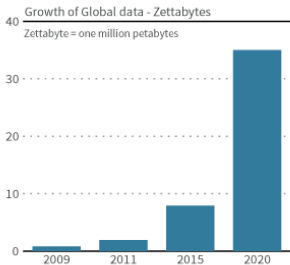
Data Revolution I

- Explosion of the capability of acquiring data on a vast scope of the human experience
- This extension of measurement ability and data collection opens immense opportunities
- In this precise moment a new product or service gets to the market as a development based on Data Science
- According to “Worldwide Big Data Technology and Services, 2012–2015 Forecast” conducted by IDC for the European Commission, data-driven technologies and services are expected to grow worldwide at a compound annual growth rate of 40% – about seven times that of the ICT market overall

Data Revolution II

Big data growth

Big data market is estimated to grow 45% annually to reach \$25 billion by 2015



*greater than

Sources: Nasscom -CRISIL GR&A analysis

Reuters graphic/Catherine Traversen 05/0/12

REUTERS

Why should Mathematics students enrol?

- Some of the most influential papers of the past 10-15 years in applied mathematics address problems in dimensionality reduction of high dimensional data sets and machine learning methods.
- The most cited paper in nonlinear PDE of Pierre Louis Lions (Fields medalist) dated 1992 is
 - ▶ MG Crandall, H Ishii, PL Lions, User's guide to viscosity solutions of second order partial differential equations, Bulletin of the American Mathematical Society, 1992with ca 4000 citations,
- while the recent paper of Emmanuel Candès, Terence Tao (Fields medalist), and Justin Romberg dated to 2006
 - ▶ EJ Candès, J Romberg, T Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Information Theory, 2006has by now more than 11500 citations (they were 9000 in May 2016).

Cool Maths without fear for missing cool jobs!

- You will learn latest developments in mathematical representation of high-dimensional data and mathematical machine learning (advanced tools of probability and mathematical statistics)
- You have the opportunity to learn technical aspects by attending courses offered by a similar program at DISI.
- We are also in the process of
 - ▶ establish contact with similar curricula around Europe for Erasmus exchange programs
 - ▶ organizing side activities such as Hacking Days and a Winter School

Caratterizzanti courses

- Geometry and Topology for Data Analysis
- Fourier Analysis
- Mathematical Biology
- Scientific Computing
- Stochastic Processes

Affini courses (at least 18CFU)

- Advanced Statistical Methods
- Bayesian Statistics
- Statistical Models
- Statistics of Stochastic Processes

Suggested courses in Affini and Liberi (at least 12+12CFU)

- Big Data and Social Networks
- Deep Learning
- Data Analysis and Exploration
- Data Mining
- Machine Learning
- Numerical Methods for PDE
- Research based Business Development
- Scientific Programming
- Stochastic Differential Equations
- Tensor Decomposition for Big Data analysis

Questions related to the Study Plan of MDS

Should you need any clarification about the Study Plan of MDS
please contact:



Claudio Agostinelli
email: claudio.agostinelli@unitn.it
skype: [claudio.agostinelli](https://www.skype.com/people/claudio.agostinelli)