UNIVERSITY
OF TRENTO - Italy

Know
dive

Studies on Human Behaviour

Course Introduction. Technical bits.

Mattia Zeni

University of Trento

About me

- ~~Background in Telecommunications Engineering~~
- Switched to **Computer Science** for a PhD on Human Behaviour
- Researcher in university and a couple of research institutions
- Switched to industry as a **researcher**
- **Data scientist** for a multinational company
- Working on a data collection infrastructure and how to extract insights from big data to take better business decisions

There is a common element in this: **data science**

Technical bits

- My part of the course is **practical**, with **tangible examples** and **real data** to solve **real problems**
- I would like to teach you how I **deal with a data science problem** in my day-to-day job
- I will introduce some of the **tools** and **best practices** I have learned
- Hopefully, all this will help you in your first data science job

NB: considering the mixed background of the class, I won't do hardcore programming, just basic stuff that everyone can handle.

Calendar

Date	Id	Modality	Starts at	Zoom	Recordings	Material	Content of Lesson	Professor(s)
8 Oct, 2020	L6	Async	-	-	Video	Slides	Methods and tools for data collection. Passive data: smartphones or other devices.	M. Zeni
12 Oct, 2020	-	-	-	-	-	-	i-Log data collection starts.	-
22 Oct, 2020	L10	Async	-	-	Video	Slides	Instruments for data collection. Capture in-the-time data (diachronic and streaming data).	M. Zeni
25 Oct, 2020	-	-	-	-	-	-	i-Log data collection ends.	-
12 Nov, 2020	L16	Async	-	-	Video	Slides	Data platforms for big data.	M. Zeni
16 Nov, 2020	-	-	-	-	-	-	Data from the i-Log data collection experiment is made available to the students.	-
19 Nov, 2020	L18	Async	-	-	Video	Slides	Data cleaning and methods for data preparation.	M. Zeni
26 Nov, 2020	L20	Async	-	-	Video	Slides	Sensor data analysis.	M. Zeni

Lesson 6

Methods and tools for data collection. Passive data: smartphones or other devices.

- Types of data for studying human behaviour
- Which types of sensor data
- Where to collect data from?
- Laboratory setting vs in the wild concept
- How to install the i-Log application

Lesson 10

Instruments for data collection. Capture in-the-time data (diachronic and streaming data).

- The i-Log application: introduction
  - Sensor data
  - Time diaries
- Application architecture
- Performance trade-offs
- Generated data
- How to install and use it

Lesson 16

Data platforms for big data.

- The StreamBase system:
  - What is it?
  - How it works?
  - Components
  - Technologies
  - Data model
  - Lessons learned
- Introduction to some tools: Apache Cassandra, Apache Spark, etc

Lesson 18

Data cleaning and methods for data preparation

- Good data vs bad data
- Why not all data is good for our purpose
- Data cleaning: what is it?
- Data cleaning approaches
- Data preparation
- Etl pipelines

Lesson 20

Sensor data analysis.

- How to answer to a data analysis request?
  - Hypothesis
  - Understand what to do
  - Process the data
  - Generate the results
  - Generate the final report
- Example: SmartUnitn location data

i-Log data collection

- 12 – 25 October data collection
  - You will be asked to install on your mobile devices a mobile application and collect data about yourself
  - You will use such (real) data in your final examination!
  - PS: Not mandatory to use it, but if you satisfy the requirement please do it for your colleagues if not for you!
- i-Log requirements
  - Android >= 6.0
  - Some free disk space (0-5gb)
  - Wi-Fi connection at home (unless you come to the University)
- 16 November data is made available to you