HUMAN DATA ANALYTICS: INTRODUCTION

Michele Rossi
rossi@dei.unipd.it

Lab. classes
Matteo Gadaleta
gadaleta@dei.unipd.it

Lab. Assistant
Simone Friso

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

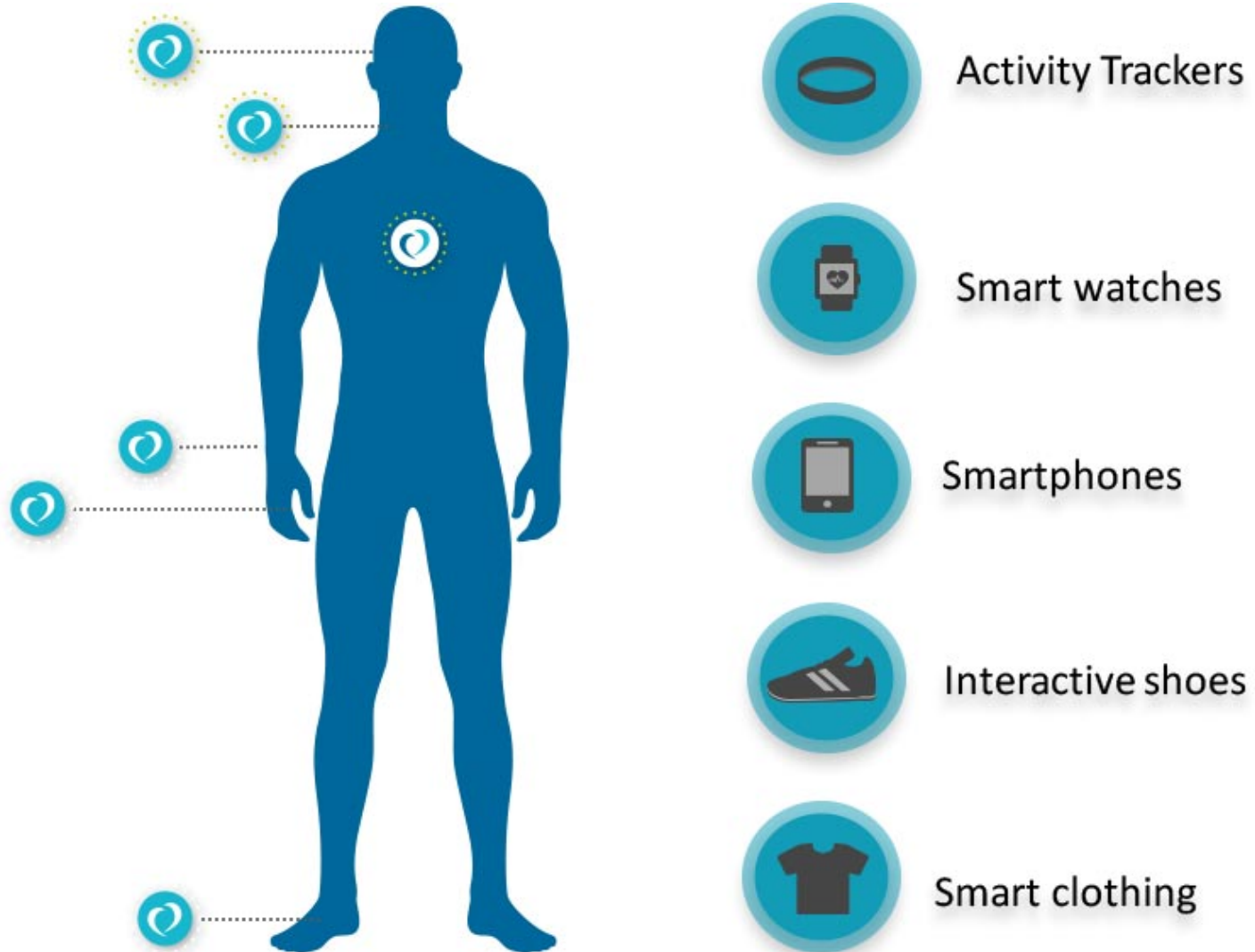UNIVERSITAS · STUDII PADUANI
MCCXXII

HDA

- A course on sensing applications
  - Centered around human-generated signals

- Master Degrees
  - Data science
  - ICT for Internet and Multimedia

- Teachers
  - Prof. Michele Rossi (theory & applications)
  - Web: http://www.dei.unipd.it/~rossi/
  - Matteo Gadaleta (lab. classes)

- Lab. Assistant
  - Simone Friso

Human sensing

# HDA in a nutshell

- Tools
  - Dimensionality reduction: PCA
  - Clustering: K-means, SOM, GNG
  - Neural networks:
    - Feed Forward (FFNN), Convolutional (CNN),
    - Recurrent Neural Networks (RNN), Autoencoders
  - Times series analysis: HMM, RNN

- Applications
  - Modeling ECG signals
  - Speech / face recognition
  - Inertial signals:
    - Authentication
   - Activity recognition (heterogeneous data)

Modeling ECG signals

- Useful for many reasons
  - Efficient TX in resource limited systems
  - Automated detection of arrhythmia
  - User identification / authentication
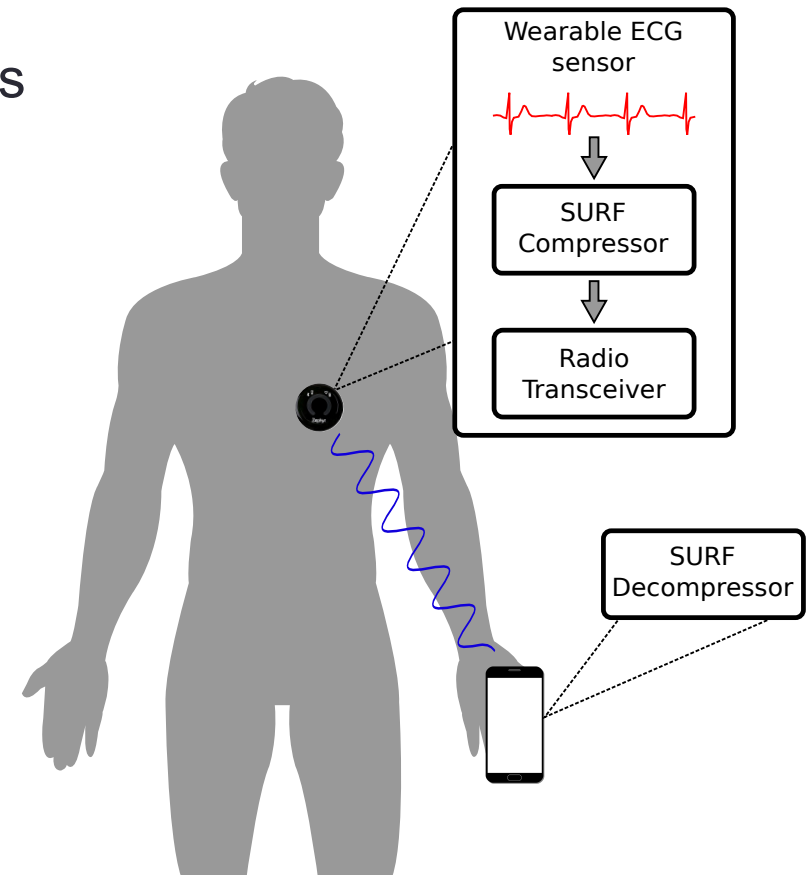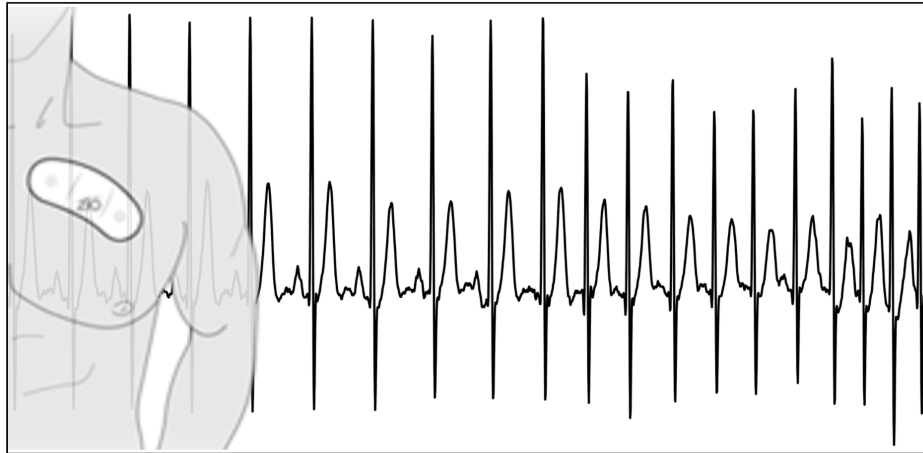
Statistical modeling of time series

- How to detect keystroke dynamics (typing behavior)
  - User identification through keystroke biometrics
- How to reliably decode words and sentences (speech)
  - Hidden Markov Models (HMM)

a voice signal example
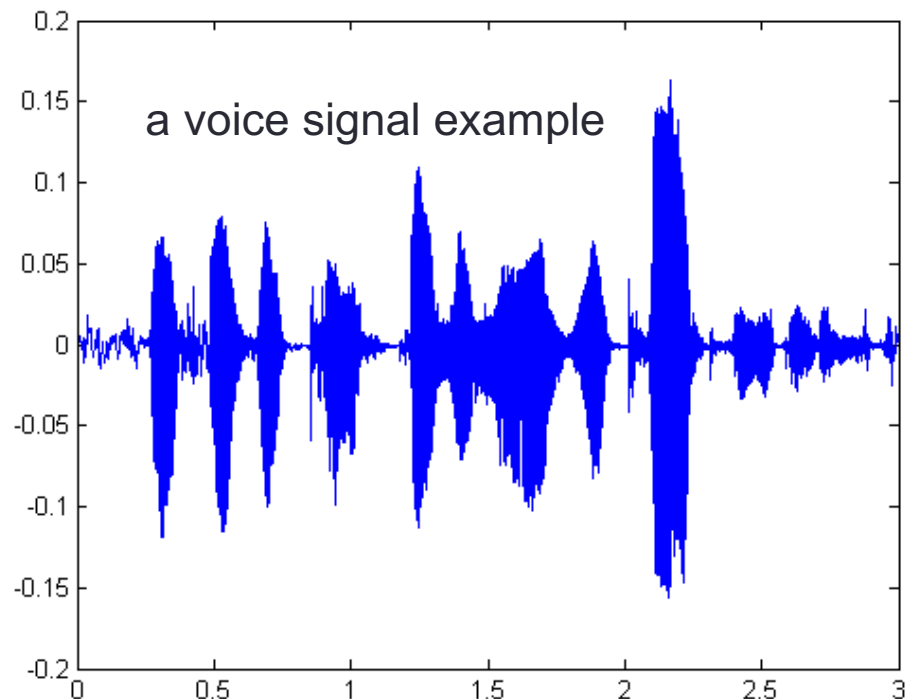
Inertial signals

- Accelerometer and Gyroscope signals
  - From Inertial Measurement Units (IMU)
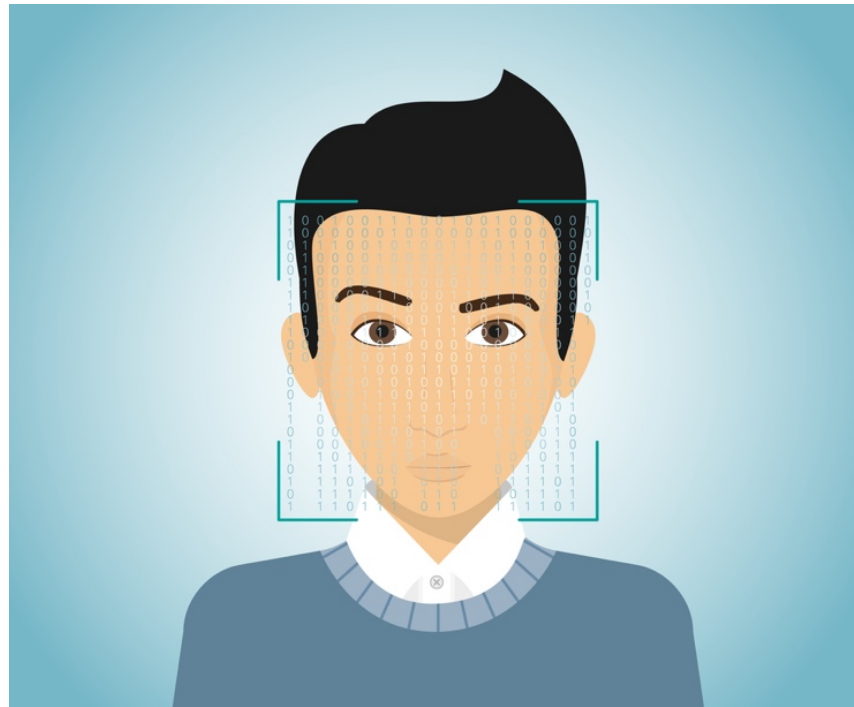   - Chest-band, wrist-band, smart watches, smart phones, etc.

- Applications
  - User authentication
  - Activity recognition

sitting standing walking running climbing stairs

Face recognition

- How to detect faces
    - Dimensionality reduction (PCA / SOM)
    - PCA / SOM + convolutional neural networks
    - (Deep) neural networks only

Course structure

1) Tools: Dimensionality reduction, clustering
   Apps: ECG

2) Tools: Hidden Markov Models (HMM)
   Apps: keystroke dynamics, speech recognition

3) Tools: Neural Networks (FFNN, CNN, RNN)
   - ECG
   - Authentication (inertial signal)
   - Face recognition

A remark

- Usually
  - Focus of basic machine learning courses is on i.i.d. data samples
  - Problems are typically: regression, classification

- This course
  - Is often concerned about modeling complex data sequences
  - Some (spatio-temporal) correlation exists among data points

- Our focus is on pattern analysis for
  - Correlated data (space, time)
  - To build applications

Course material

- Course material
  - http://www.dei.unipd.it/~rossi/courses/HumanData/HDA.html

  User: HDA-student

  Passwd: Hda_0q1o9w2i_MR;

- Mailing list
  - hda@dei.unipd.it

Exam

- Project based
  - Project assignment from instructors
  - Use a machine learning algo on a public dataset

- Outcome
  - Written project report (max. 15 pages)
  - Power Point presentation (running code is appreciated)

- Groups
  - 2 students per group

Guidelines about: public database, task to be performed and project report structure will be provided when we start our laboratory activity

Lab. Classes (1/2)

- Will feature:
  - An introductory lesson on:

  - Guided coding sessions to build machine learning apps
    - Build everything from scratch (simple applications)
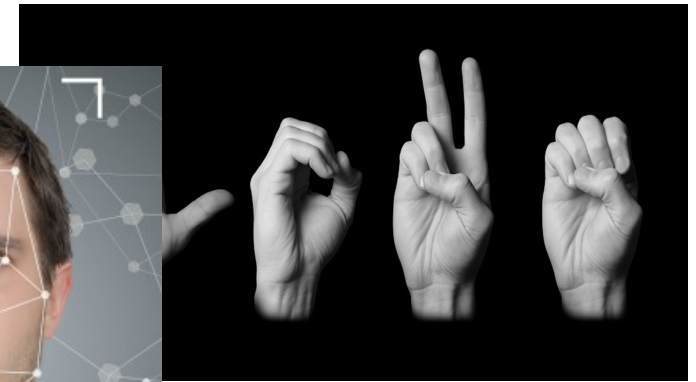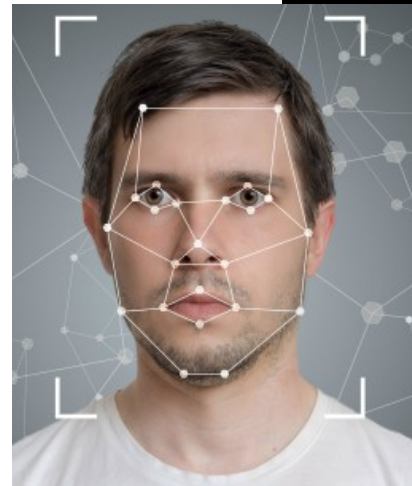   - Use some popular libraries (more complex projects)

Lab. Classes (2/2)

- Topics and Applications
  - Logistic regression for classification
  - Shallow/Deep Neural Networks for image classification
    - Cat classifier

  - Convolutional Neural Networks
    - Face recognition
    - Sign language translation

  - Object detection and recognition
    - Car/signs/pedestrian detection for self driving cars

HUMAN DATA ANALYTICS: INTRODUCTION

Michele Rossi
rossi@dei.unipd.it

Lab. classes
Matteo Gadaleta
gadaleta@dei.unipd.it

Lab. Assistant
Simone Friso

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE