Data Mining

Livio Finos

My self

- Livio Finos livio.finos@unipd.it

- Prof Statistics - Dept. Statistical Science

- I like: Multivariate analysis, Statistical Inference

- Applications: Neuroscience, Ecology, Psychology, Genetics, Text-mining
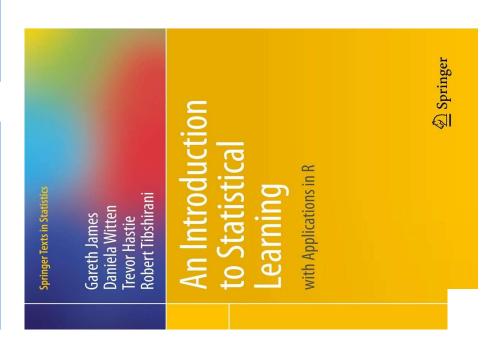
- and *Who are you?*

Text Book

An Introduction to Statistical Learning - with Applications in R

Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani

Download it here or here (Python codes too)

Springer Texts in Statistics

Gareth James
Daniela Witten
Trevor Hastie
Robert Tibshirani

An Introduction
to Statistical
Learning

with Applications in R

Springer

Software

- R (+ Rstudio?) The didactic material is written in R (and I release most of it)

- Any other software (e.g. Python, Matlab) you like/know is accepted (but I can't provide much support in this case)

- Bring your laptop with you in class (at least in pairs of studentes) so that we can touch the data and the methods

Activities + Exam

You choose (this week) between:

· Team Based Learning (TBL) modality

· or Standard one

Team Based Learning

- You work in Groups (around 6 people)

- 3 Activities along the semester

  - Long-lasting prediction competition Phone Users (max 4 pt.s)

  - TBL1: 3-4 lessons lasting activity (from 27/03?) (max 3 pt.s)

  - TBL2: 3-4 lessons lasting activity (from 22/04?) (max 3 pt.s)

- Final exam (individual activity): Laboratory (practice with some question about the theory) (max 22 pt.s)

- Groups – formed by me – last throughout the course

Other students (no TBL, Standard mode)

- You work alone
- Prediction competition Phone Users + small project + oral discussion (max 10 pt.s)
- Final exam: Laboratory (practice with some question about the theory) (max 22 pt.s)

I have a strong preference for the TBL mode: you are motivated to work along the course and you get less stressed for the final exam.
But your are free to choose, of course.

How Long-lasting Competition works

- Prediction competition Phone Users

- Partial Scores before 10 June,

- Closure at 10 June,

- Evaluation is based on Final Score (50%) and Quality of the Report (50%)

How TBL works

- Phase 1: Independent study (not in classroom)

- Phase 2: Individual RATe (iRAT); Day 1, 15 mins

- Phase 3: Team RATe (tRAT); Day 1, 25 mins

- Phase 4: Burning questions; Day 1

- Phase 5: Complaint; on moodle forum, after Day 1

- Phase 6: Lesson/lab based on your Burning questions; Day 2

- Phase 7: Team APPlication (tAPP); Day 3 + HomeWork

- Phase 8: Peer Evaluation (online + briefly in class)

Evalutation of TBL

- iRAT: 30%

- tRAT: 30%

- tAPP: 30%: evaluation from students (with my veto)

- peer evaluation (within group): 10%