



Visual Analytics

X– Final Assignment

Exam

Exam is composed by 2 main parts

- **paper discussion**
- **final assignment**

Paper discussion

- Assignment of a research paper to a group of students (minimum 2, maximum 3 students)
- Analysis of the paper and discussion (with a slides presentation) of the paper to the room
- Each student MUST explain a part of the paper
- Paper proposed by us (will be available on the official website of the course)
 - bidding following a First-come-First-serve rule (with final tuning by us)
- To be discussed in the final part of the course

Rules

Final assignment can be completed **individually** or in **group** of MAX 3 persons.

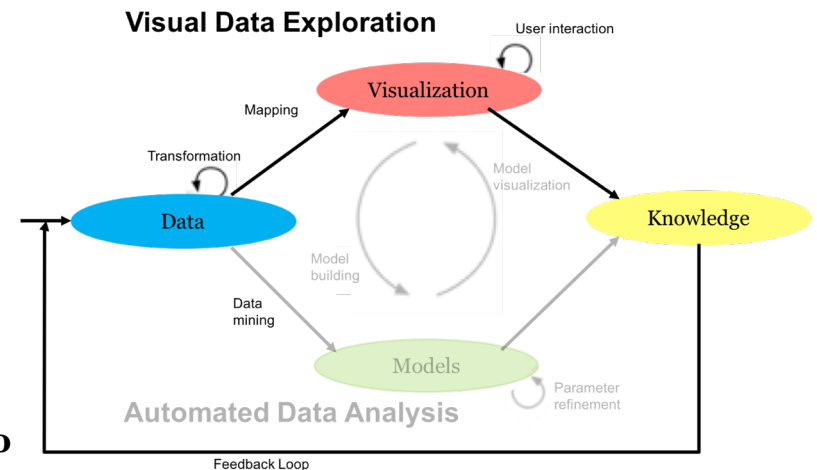
- If, for any reason, you did not discuss the chosen paper during the class you have to present it together with the final assignment

2 tiers of projects:

- **Project theme proposed by the professor:**
 - We will provide specification for the project (dataset, goals, eventual additional requirements) on the website
- **Free proposal**
 - Project proposal comes from the students group

Rules

- **Preliminary approval needed:** before starting working on the assignment you MUST send us (by e-mail to Prof. Santucci with Dr. Angelini in cc) a 1 page draft of the idea that:
 - Specify the general idea (Analytics part, Visual Part)
 - Specify the dataset (characteristics and context)
 - Specify the various characteristics in Relation to the Visual Analytics cycle



Characteristics

- **Mandatory:** every assignment **MUST** use a PCA or a MDS or a t-SNE algorithm in order to:
 - **Preprocess** the dataset for choosing the right dimensions
OR
 - Having it **integrated** as **ANALYTICS** part
- **ANALYTICS:** every assignment **MUST** contains at least 1 computation that is triggered by user interactions (no change in datasets, no simple filter)
- **VISUALIZATION:** every assignment **MUST** have a visual part constituted by at least 2 Visualizations coordinated in both ways and interactive

Datasets

- You are all encouraged to propose a context and a datasets of your choice on argument that you find interesting to analyze in the following fields:
 1. An application close to the topic of the chosen paper
 2. Sport analytics (football, tennis, sailing, etc.)
 3. Bitcoin, block chain
 4. Traces collecting user interactions along an experiment (contact Prof. Santucci, Dr. Angelini for details)
 5. Network traffic data
 6. Social network
 7. Vast 2018 challenge
 8. Data coming from the datasets described in the following
 - The Dataset **MUST** respect the rule that the index AS (AngeliniSantucci) defined as:

$$\mathbf{AS} = \mathbf{\#tuples} * \mathbf{\#dimensions}$$

is contained in the range 10000 – 50000 (and more for the braves.....)

2. Group Project

DataSets:

1. Data from ISTAT:
 - Many different datasets available form a main data-warehouse

<http://dati-censimentoindustriaeservizi.istat.it/>

2. Italian Governement Open Data

<http://www.dati.gov.it/catalog/dataset>

3. Top World Universities Rankings:

<http://www.topuniversities.com/university-rankings>

2. Group Project

DataSets:

4. **IEEE VAST-Challenge** (Conference on Visual Analytics Science and Technology)

- 3 different mini-challenge, application for 1 of them
- An e-mail address is required for downloading the dataset
- Possibility for a publication in the proceedings if submitted before the deadline (July 13th)
- Interested people have to coordinate with us

<http://www.vacommunity.org/VAST+Challenge+2018>

5. Users Ratings on Movies:

- Also IMDB can be a good resources for well-known ones.

<http://grouplens.org/datasets/movielens/>

6. UCI datasets:

<https://archive.ics.uci.edu/ml/datasets.html>

2. Group Project

DataSets:

7. Italian Top Scientists:

http://www.topitalianscientists.org/top_italian_scientists.aspx

8. Kaggle datasets:

<https://www.kaggle.com/datasets>

9. Cyber-Security datasets:

<http://vizsec.org/data/>

10. Network data:

<http://snap.stanford.edu/>

- It is possible to propose own datasets obtained by databases, projects, scraping ...

Contents and time

- Final assignment **MUST** contains:
 - The running **software**
 - A relation (5-6 pages) **similar to a paper** that describes the whole design process, rationale and prototype, comprehensive of references
 - A slide deck **presentation**
- It must be discussed during the exam date with a time of 20 minutes + 10 minutes of Q&A and discussion

Evaluation

- The final grade is the result of this formula:

$$\text{final grade} = 1/3 (\text{paper discussion}) + 2/3 (\text{final assignment discussion})$$

- Laude will be assigned for particularly brilliant characteristics and at commission's discretion

Dates

- Exam dates are the normal dates reported in Infostud
- **You must make a reservation through Infostud to attend the exam**
- In order to discuss the final assignment the students **MUST** have already presented the paper
- Both of these activities (paper presentation + final assignment presentation) can be done during the same exam date and in the same day