

## FINAL REPORT

### Eu-SPRI Summer School

# Tools and methods for analysing complex Science, Technology and Innovation (STI) systems: A gentle introduction to Machine Learning (ML), Network Science (NS) and Text Mining (TM)

IRCrES-CNR, Rome – July 1-3, 2019

### Chairs of the conference

Antonio Zinilli

Giovanni Cerulli

### TEACHING STAFF

Dr. Antonio Zinilli (IRCrES-CNR), Dr. Giovanni Cerulli (IRCrES-CNR), Dr. Carlo Drago (Università Cusano)

### Local Organising Committee

Antonio Zinilli

Giovanni Cerulli

Marco De Biase

Emanuela Reale

## Participants

33 researchers participated to the Summer School:

<b>USPRI Members</b>	5
<b>No USPRI Members</b>	26
<b>Researchers from the CNR</b>	2

## OBJECTIVES OF THE COURSE

Machine Learning (ML), Network Science (NS) and Text Mining (TM) are relatively new techniques able to enlarge our understanding of complex socio-technological systems, either by digging deeply into the data informative power (ML and TM), or by increasing the understanding of the system relational dimension (NS). The training provided participants with the essential tools for a correct application of some popular ML, NS, and TM methods in various STI contexts. In particular:

- ML techniques proves useful for factor importance detection, as well as for classification purposes in a model-free stance;
- NS techniques are useful to identify and study structure and dynamics of large and complex STI communities;
- TM is a set of tools able to extract relevant information from articulated documents, both for descriptive and policy analysis purposes.

The course provided three modules (one on ML, one on NS, and one on TM) with the aim of balancing theory and applications. Participants carried out some exercises assigned by the instructors under their supervision.

## Social dinner

On September 2<sup>th</sup>, at 7.30<sup>PM</sup>, the social dinner took place at "Efeso il Barrocciaio". Besides being another opportunity for networking, this dinner offered Summer School participants the opportunity to taste the typical Italian food.

## Evaluation of the conference

Overall, the attendees were highly satisfied with the program and activities which is confirmed by the attendee evaluations. Attendees graded the overall conference with a good evaluation judgment. High scores were particularly given for the overall atmosphere, the feedback received and scientific program. Staff members were also really enthusiastic about the group and enjoyed the interactions with the young researchers.

## Summer school program

### Tools and methods for analysing complex Science, Technology and Innovation (STI) systems: A gentle introduction to Machine Learning (ML), Network Science (NS) and Text Mining (TM)

IRCrES-CNR, Rome – July 1-3, 2019

---

#### Day 1 – July 1, 2019

(Antonio Zinilli)

9:30

Registration

9:45-10:00

Welcome

10:00-11:00

Basic concepts of Social network analysis

11:00-11:30

Coffee break

11:30-12:30

Main Network models (ERGM)

12:30-13:30

Lunch

13:30-14:30

Application scenarios (some illustrative examples of Network Science  
on specific datasets of Science Technology and Innovation (STI) systems)

14:30-15:30

Organization of the laboratory assignments  
Creation of groups and provision of data

15:30-16:00

Coffee break

16.00-17.00

Young participants' presentation of their assignments

---

## Day 2 – July 2, 2019

(Giovanni Cerulli)

10:00-11:30

An introduction to Machine Learning and Data Science for analyzing complex STI systems:  
Identification, prediction, trade-offs, and validation approaches

11:30-12:00

Coffee break

12:00 13:00

Monte Carlo, Bootstrap, and model Validation: a primer

13:00-14:00

Lunch

14:00-15:00

Model selection and regularization:  
Optimal subset selection; Shrinkage Methods: Lasso, Ridge, and Elastic regression

15:00 – 16:00

Laboratory with Stata and R

16:00 – 16:30

Coffee break

16.30 – 17.30

Applications to real STI datasets

19.30

Social Dinner

---

### Dinner on July 2, 2019

19:30 Meeting at Via dei Taurini, 19

---

## Day 3 – July 3, 2019

(Carlo Drago)

10:00-11:00

Concepts and tools for Text Mining analysis

11:00-11:30

Coffee break

11:30-12:30

Text Mining applications with R

---

## Pictures from the Conference



