



DOCTORAL
SCHOOL
AT GDAŃSK
UNIVERSITY
OF TECHNOLOGY

Course: Scientific methods of computer data analysis and presentation

Teaching hours: 30 h

Prerequisites: The course is primarily open to all PhD students at Gdansk University of Technology.

This course is compulsory for PhD students assigned to Technical Informatics and Telecommunications tracks at Doctoral School at Gdańsk University of Technology

Course outline

Content

This module is all about getting the student to be prepared to analyse and synthesize results of scientific experiments. The sessions provide essential information that you require to report the experimental results to a broader audience, especially in scientific journals and presentations. The course is designed to deliver to students knowledge on essential statistical methods in experimental data reporting, design of experiments and appropriate presentation of the input data and results. Throughout the course the students should gain skills to: parse experimental results and other ancillary data; conduct statistical tests, use statistical models and sample data; use basic supervised and unsupervised learning methods; report the data using variety of figures or tables. The course will be conducted using the R language with basic and additional packages, especially ggplot2. Students are free to use other comparable tools such as python with plotnine. The lectures will show the background of the methods used in the laboratories; best practices in scientific data reporting; and also will provide skills for critical judgments of other experimental results. During the seminar students will show and discuss their experimental data elaborated using the techniques presented during the course.

General topics coverage:

1. Data parsing .
2. Data sampling.
3. Statistical tests and models.
4. Regression and classification.
5. Various figures and proper plotting.
6. Seminar.

Teaching mode

There will be 20 hours of laboratories, 6 hours of lectures and 4 hours of seminars, to be completed during the third semesters of PhD programme. The teaching method is basically computer laboratory using the R software and various data combined with seminar discussion. During the course students will be asked to introduce themselves to the basic R functions (see literature). The course is entirely delivered in English.

Examination

The PhD students will be required to demonstrate their skills, knowledge and understanding of scientific methods of computer data analysis and presentation during an written examination. Students will also have to presents, discuss and defend the results of their data elaboration during the seminar. Before examination the students are requested to pass the laboratory classes.

Fundamental readings:

1. Wilkinson, Leland. "The grammar of graphics." Handbook of Computational Statistics. Springer, Berlin, Heidelberg, 2012. 375-414. <https://link.springer.com/book/10.1007/0-387-28695-0#toc>
2. Tufte, Edward R. The visual display of quantitative information. Vol. 2. Cheshire, CT: Graphics press, 2001.
3. W. N. Venables, D. M. Smith and the R Core Team. An Introduction to R. <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>
4. Łukasz Komsta. Wprowadzenie do środowiska R <https://cran.r-project.org/doc/contrib/Komsta-Wprowadzenie.pdf>
5. ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics <https://cran.r-project.org/web/packages/ggplot2/index.html>