Theoretical Foundations of the Analysis of Large Data Sets

The course will be based on Lectures of Professor Emmanuel J. Candes from Stanford University available at

https://statweb.stanford.edu/~candes/teaching/stats300c/Lectures/Lectures2018.zip

https://statweb.stanford.edu/~candes/teaching/stats300c/lectures.html

Description (by Prof. Candes) : The main goal of this course is to expose students to modern ideas in statistical theory. Whereas classical theory is concerned with the behavior of statistical estimates when the number of variables is fixed and the sample size increases, our emphasis is on statistical inference in high-dimensional settings where there may be as many, or more, variables than observations. Our focus is motivated by always newer technologies, which now produce extremely large datasets, often with huge numbers of measurements on each of a comparatively small number of experimental units.

The necessary conditions for getting positive grade for the course are

- ***the positive grade for the Lab course;***
- ***the positive points for the 1st and 2nd mid-term exams.***

Course grade will consist of:

1) 1st mid-term exam (will be given on 17.11) – 35 points (max)
2) 2nd mid-term exam (will be given on 12.01) – 35 points (max)
3) Lecture additional points – 10 points (max)
4) Points for labs – 25 points (max)

and corresponds to the following

points	grade
>=90	5
>=80	4+
>=70	4
>=60	3+

Students who do not pass the course based on labs and mid-term exams will be given a chance to write an exam during an exam session. The tentative date for the exam is … of February.

Lab course grade will consist of points for each Lab list (5 lists x 5 points= 25 points (max)) and corresponds to the following

points	grade
>=22,5	5
>=20	4+
>=17.5	4
>=15	3+
>=12,5	3

Points for single report:

0 point. - no report
<2 – incomplete solution of practical part (less than 50%);
2 points – incomplete solution of practical part (more than 50%), no comments;
3 points – the solution of the practical part is almost complete (more than 90%), but comments are not aesthetic or in-depth, the theoretical part may be incomplete.
4 points – complete solution of the practical part and more than 50% of the theoretical part, the solution is extended with own comments;
5 points. - a complete solution with in-depth reflections and comments, testifying to the student's in-depth understanding of the matter to which the task was related.

The necessary condition for getting positive grade for the Lab course is to obtain 2 or more points for each of the labs.