Lab: Statistical Inference – DEIS DT8007

Supervisor: Martin 2019/9/12 (~ 2 hours)

Today you will do something hands-on with pattern recognition and statistical evaluation!
Goal: to have an idea of how to

- set up a system to solve simple classification problems
- evaluate simple properties of a system using statistics

Approach: you will require

- Facilities. You should have access to these instructions, and access to some computer with a way to program and statistics software like Excel. You can work in your groups, but you need to state what each person did.
- A short pdf report, to be submitted by email before October 12 to Martin (marcoo@hh.se), including code, calculations, and/or conclusions. Any feedback will be given after the deadline.

Result: Your report will be judged as passing, (reasonable but) failing, or unsubmitted.
Please note:

- If the report is insufficient, you will get a chance to review and resubmit.
- All labs must be submitted to receive a passing grade in the course.
- The filename of the report should be something like "deis_lab1_group1.pdf".
- **You are expected to do original work throughout this course. This applies also to labs. Using someone else's code/report/etc without referencing them is a serious problem which you can expect to result in a failing grade for the course (and possibly more). We don't have any tolerance for this.**

Your tasks:

1. Pattern recognition

Please carry out some pattern recognition task, using some data/recognition strategy you find interesting, on something which can be related to your project (robots, terra-forming, cyborgs, etc). The important thing is for you to try to learn.
If you want a little programming experience, a nice idea could be to code your own algorithm (e.g. k-NN, k-means, perceptron), make yourself a little dataset (e.g., write some numbers into a text file, or get some numbers from your robot's sensors), and classify.
Or, if you want some experience working with existing tools, you can download some library, dataset, or code (e.g., for deep learning) to modify; just make sure to reference it and be clear what you did/what you did not do; e.g. you can find some datasets at:
https://towardsdatascience.com/top-sources-for-machine-learning-datasets-bb6d0dc3378b
Briefly describe decisions you make (features, parameters, etc) and results.

2. Statistical evaluations.

Please carry out some statistical evaluation task, using some data/approach you find interesting, on something which could be related to your project (robots, terra-forming, cyborgs, etc).
You can download data (with a reference in your report) or make your own data.
Please briefly describe: what is your hypothesis, what is the null hypothesis, what is alpha, what statistical hypothesis test you picked and why, and what is the result?