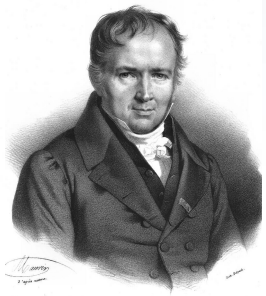
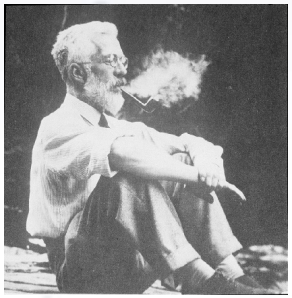
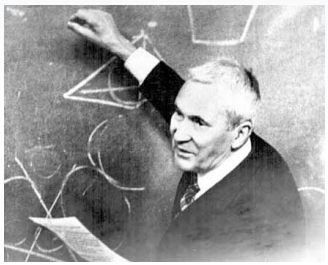
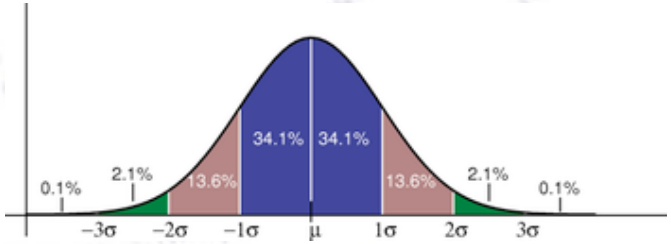
Big Data Analysis

Final Project Exam

Troels C. Petersen (NBI)



"Statistics is merely a quantisation of common sense - Machine Learning is a sharpening of it!"

Your presentations

Each presentation is allotted 15 minutes, sharp!

Please make sure, that you note the time of starting, and we will try to let you know, when the time limit is closing in.

We like questions, but in the interest of saving time, please think about them ahead of time, and state them in a short and accurate fashion.

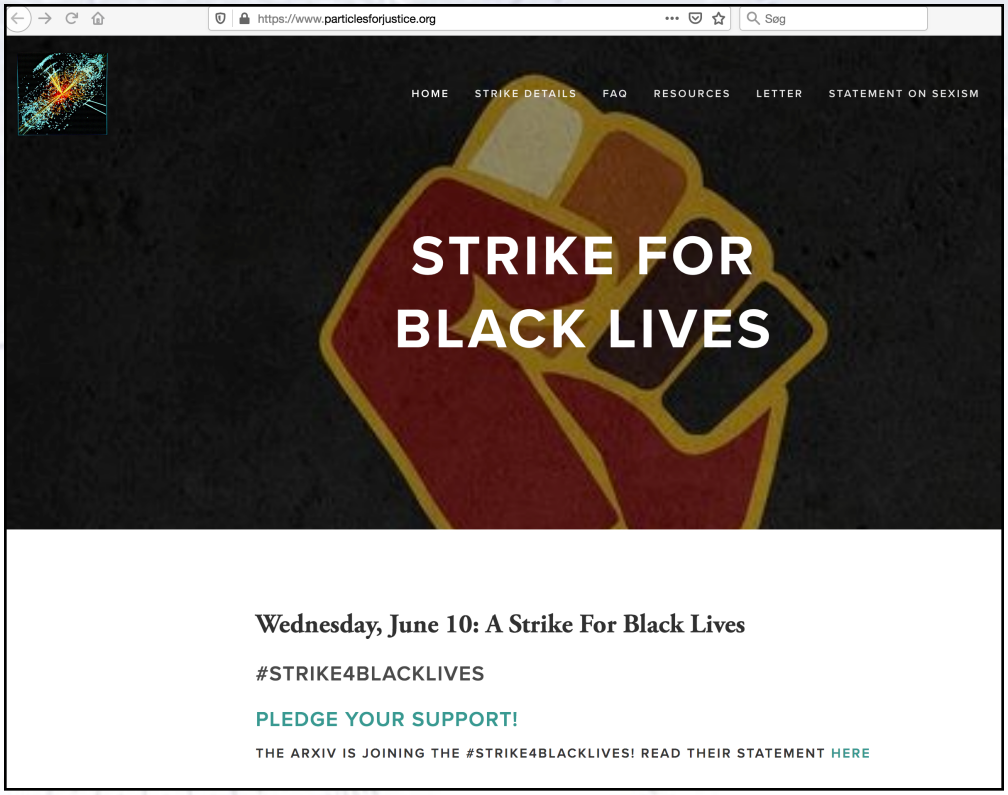
Be ready to start presenting, when you are up next.

The final projects will be evaluated based on the following criteria:

- Complexity of problem and depth of solution (incl. appendix)
- Choice of methods and arguments behind
- ML performance and own evaluation of it
- Clarity of presentation
- Implementation, technical details, optimisation, etc. (your appendix)
- Ability to evaluate ML usage (your evaluations of other presentations)

Go on strike? Cancel exam?

Some of our (mainly US) colleagues called for a strike in academia… today!



We of course fully agree that black lives matter, and that racial police brutality should not be tolerated in any way.

This came up with too short a notice to reschedule, and so to honor your work, we decided not to cancel the exam. However, we do take the matter seriously.

Start:	Group member names:	Project Name:
9:00	Troels C. Petersen	Short introduction
9:15	Rasmus FØ, Peter C	Reconstructing neutrino events in IceCube
9:35	Maria, Mads, Andy, Emil	Walmart data (from Kaggle)
9:55	Katja, Helena, Viktoria, Simon	Skin lesion dataset (from Harvard)
10:15	Break	
10:45	Marta, Ann-Sofie, Emy, Yanet	Retrieval of sea surface temperatures
11:05	Christopher, Nikolaj, Joakim	Predicting super conductors critical temperatures
11:25	Aske, Mikkel RS, Anna, Mikkel LL, Moust	Reconstructing neutrino events in IceCube
11:45	Lunch break (1 hour)	
12:45	Rasmus MS, Haider	Estimating publication year of music
13:05	Edwin, Miren, Alba, Fynn	Blood cell type data (from Kaggle)
13:25	Runi, Simone, Marcus, Jonathan	Calibration of new astro-data for exo-planet
13:45	Break	
14:15	Sofus, David, Elias, Kristoffer	Learn to trick face-tracking systems
14:35	Dina, Aline, Albert, Michael	Wheat data (from Kaggle)
14:55	Laurent, Orestis, Griogos, Carlos	Extraction of sentiment from Tweets
15:15	Break	
15:45	Svend, Julius	Estimating genre of music
16:05	Emil Schou Martiny	Own very noisy data and AstroData?
16:25	Nicolas Remy Høegh Pedersen	Identification of objects in 2D images
16:45	Break	
17:00	*Unforseen events/projects/summary*	*(Hopefully not needed!)*

Bonus Slides