Goal	
Be able to predict how long a patient is going to stay in ICU at Beth Israel Deaconess Medical Center	

Method	
Theory	1. CITI ethics course 2. Simple linear and multivariable regression, logistic regression 3. One-hot-encoding and entity embedding, semantics
Tools	Python, Neural Networks (maybe tensorflow)
Data examples	MIMIC-III Clinical Database

Literature
1. Jay L. Devore & Kenneth N. Berk: Modern Mathematical Statistics with Applications, second edition, 2012. Springer. ISBN: 978-1-4614-0390-6. 2. https://towardsdatascience.com/introduction-to-linear-regression-in-python-c12a072bedf0 3. https://www.fast.ai/2018/04/29/categorical-embeddings/ 4. https://arxiv.org/pdf/1604.06737v1.pdf

Time schedule
1.may - 22.july

Guidance
Email, whenever

Milestones	
Pass ethic course and (get access to MIMIC data set - 28.may)	4.may
Readings on regression and one-hot-encoding	26.may
Data extraction and integration from MIMIC data , extracting relevant classification/domain information from disease ontology.	5.june
Training neural network on one-hot encoded data for predicting no. of stays in ICU (changing number of hidden layers, activation functions, learning rate etc and finding the best model)	13.june
Encoding data categorical data using disease domain information and training neural networks (that's the challenging part and new thing to do)	14.july
Writing paper	21.july