* Training listing(?) phänomen und wie verhindert man es?
* Netzwerk mit 1 Hidden Layer
* Regression und ForwardProp
* Derivative Hyperbolic Tangent Show
* Backpropagation, derive the camps accuracy to weights in the last layer of gradient of loss function
* Implement Method for gradient desc, for given weights und training points
* Def gradient\_desc()
* CNN given
* Name 4 Valid Input dim using zero passing (w,H,Channel)
* Calculate Number of trainable parameters for each layer
* Str derd Convol + Max Pool
* If we remove Max Pool and Setstride of 2D Conv, how would output of flatten change?
* Explain 2D Max Pool (Sketch)

MC:

* Role of Bias
* Softmatrix actv Function
* Least suitable output neuron <-> lost function pairing
* Dying grads and which act function?
* J =-E yilog(yi) ?
* Momemntum opt. alg -> acc. Convergence?
* Alg. Problem: Monotonically desc. Learning
* Learningrate of first + second moment of grad, which alg?
* -Mode Collapse
* Minibatch Learning, rponh has n advantage of shuffling n? n=?
* Grad desc most likely to converge?
* Ensemble method?
* Dropoutrate p= how man sep. Net?
* Compressing NN
* Why less overfit CNN
* -back prop Max pull
* RNN Current State form
* GRU Gate State
* Vanish Exp GPTT + RNN
* S2S Learning
* CTC vorteil
* RI Schema
* Q learning
* Deep Q
* GAN