

Introduction to Deep Learning

MLP, gradient descent backpropagation

Andreas Krug, M.Sc.

ankrug@uni-potsdam.de



22. October 2018



Reading assignments





How well did you understand the reading topics? (either from the DL book or other material)

less

more



Explain the back-propagation algorithm and its importance for neural network training!



How does gradient descent work? Bonus: What are the differences between stochastic, batch and mini-batch gradient descent?



Why do we need non-linearities?



Which activation functions do you know? Sketch them and discuss their pros and cons!



How do the choice of output units and the loss function relate?



Why is the negative log-likelihood a popular choice as loss function?



What is Maxout? Sketch a layer with Maxout activation!



Why use a linear hidden layer?



What does the universal approximation theorem (practically) mean?



Sketch and explain the computation graph for the cross-entropy loss of an MLP with 1 hidden layer!



Debriefing

- Point 9 is most import finding
- Apart from computing the derivative of the cost function, what are other applications of the backpropagation algorithm in the context of deep learning?
 - optimize input -> analyze the learned model (introspection)



- Join Campus.UP Workspace "Introduction to Deep Learning WS2018"
 - introduce yourself in a short blog post including Picture, Name, Background, Motivation
- Join Mattermost
- Reading: Convolutional Neural Networks (CNNs)
- Programming exercise: using visualization to debug DL code

 Slides & assignments on: <u>https://mlcogup.github.io/idl_ws18/</u>