Aalto-yliopisto
Perustieteiden
korkeakoulu

Gradient based optimization methods in Feedforward Neural Networks (topic presentation)

Julius Lind

14.06.2019

Advisor: Prof. *Harri Ehtamo*

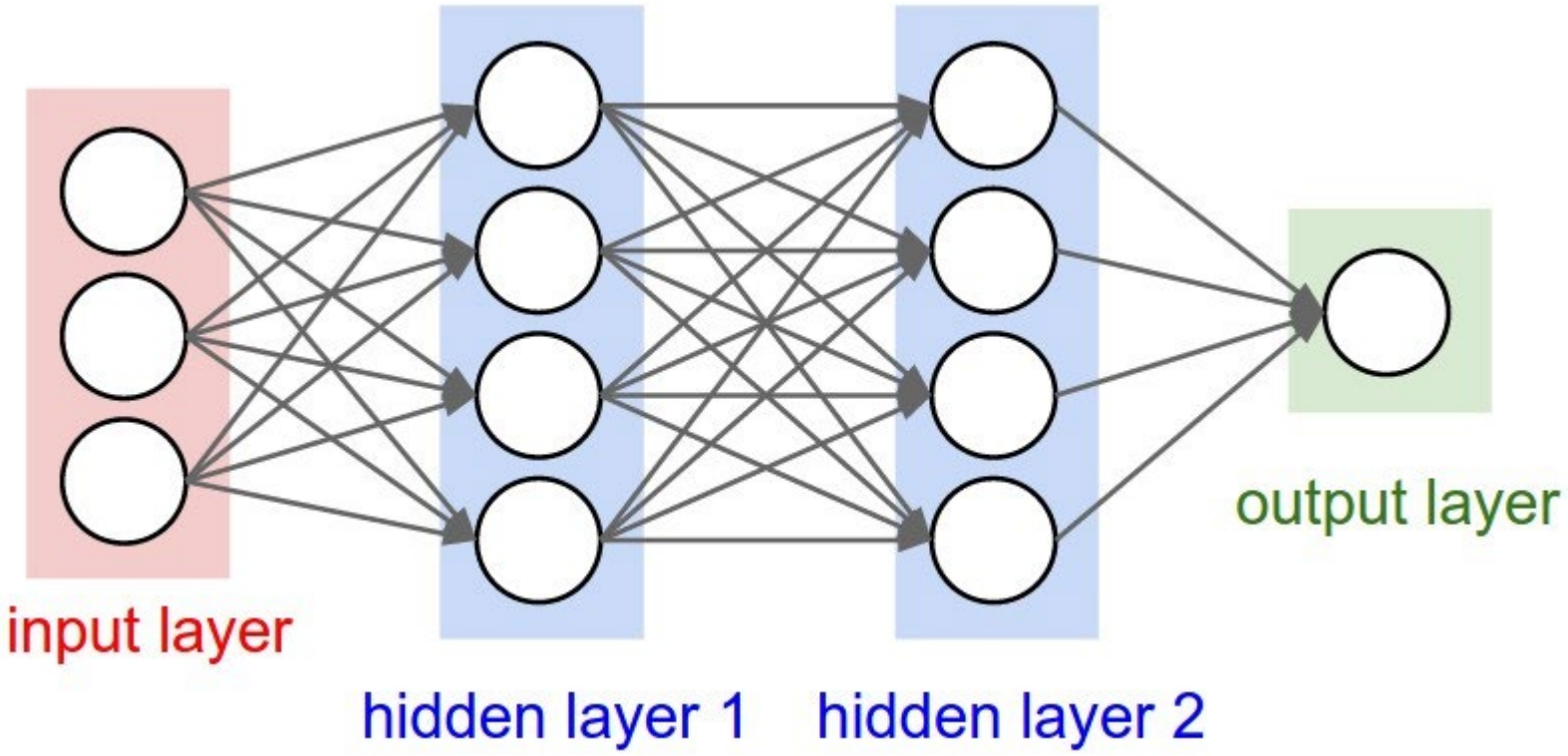
Supervisor: Prof. *Harri Ehtamo*

Systeemianalyysin
laboratorio

Background

- Feedforward neural networks are a class of Artificial Neural Networks (ANN)
- ANNs can be used for:
  - Image classification
  - Machine translation
  - Regression tasks
- Feed forward neural networks consist of an input layer, an output layer, and optional hidden layers in between
- Input data (X, y)
- Each layer applies a transformation on its inputs, and feeds it to the next layer:

$$z = \sigma(AX + b)$$



Adapted from Stanford course CS231n Convolutional Neural Networks for Visual Recognition, accessed 01.06.2019

Background

- Output z compared to the target output y to compute loss:
  - MSE = $\frac{1}{n}\sum_{i=0}^{n}(y_i - z_i)^2$
- Optimization problem:
  - Minimize loss wrt. A,b
- Backpropagation
  - Use chain rule to calculate the gradient of the loss wrt. A,b
- Gradient can be used in gradient based methods to minimize loss
- Problems:
  - Large number of optimizable parameters
  - Large number of data points
  - Non-convex objective functions

Objectives

- Present training of feed forward neural networks as an optimization problem
- Present different gradient based optimization methods:
  - Gradient descent
  - Stochastic gradient descent
  - Momentum
  - Adaptive learning rates
  - Adam
- Test the performance of different gradient based methods on a large scale neural network
  - Final loss
  - Training time

Tools, Data and references

- PyTorch and Scikit-learn
  - Open source machine learning libraries for python
- MNIST dataset
  - Image classification dataset consisting of 60000 images of handwritten digits

- Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016.
- Sebastian Ruder. An overview of gradient descent optimization algorithms. arXiv e-prints, art. arXiv:1609.04747, Sep 2016

Schedule

- 14.06.2019: Topic presentation
- 06/2019: Literature review
- 07/2019: Performance tests
- 08/2019: Final presentation