MNIST Training for BNN

Jack Diep, Florian Köhler, Yannick Naumann

September 3, 2021

Design Your Own CPU - Design of Embedded Systems

Content

Binarisation of Linear Layer

- binarisation of weights
- binarisation of input data
- calculation through *nn.linear*

Activation

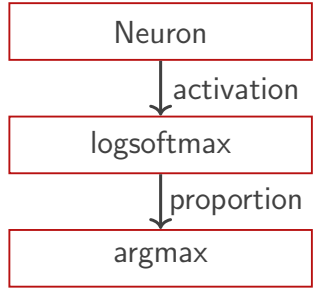
Inhalt...

Batch Norm (BN)

- In NN
  - normalize batches
  - mean 0
  - standard derivation 1
- In BNN
  - prevent *expolding gradient*

Evaluation of last layer

- normalisation of activation
- decision of the network



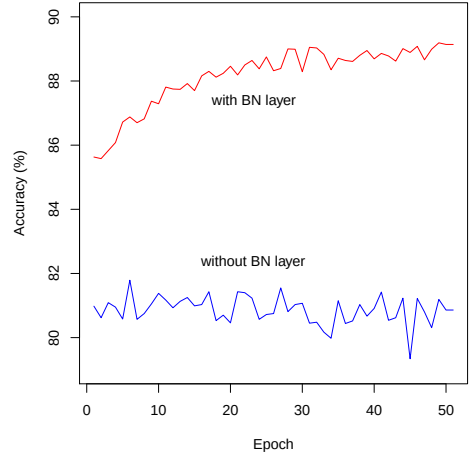
1. Neural Networks

2. BNN Design

3. **BNN Training Analysis**
   - Layer Analysis
   - Parameter Analysis

Effect of Batch Norm

- 7.4% improved peak performance
- Less jitter with BN
- Reduced expolding gradient



Consequences of linear layer binarisation

Run	binary	normal
1	**88.29**%	**97.43**%
2	87.32%	96.98%
3	87.19%	97.2%

- training for 50 epochs
- mean loss of 9,6%
- loss in granularity

learning rate

- higher value $\rightarrow$ more weights are updated
- balance between over- and underfitting

evaluation learning rate