

A Neural Network Toolkit for MT

Maximiliana Behnke
Tomasz Dwojak
Marcin Junczys-Dowmunt
Roman Grundkiewicz
Andre Martins
Hieu Hoang
Lane Schwartz

A Neural Network Toolkit ~~for MT~~

Maximiliana Behnke
Tomasz Dwojak
Marcin Junczys-Dowmunt
Roman Grundkiewicz
Andre Martins
Hieu Hoang
Lane Schwartz

Why create another NN toolkit?

- Flexibility
 - Add functionality easier & faster
- Speed
 - Pure C++ implementation
 - GPU-enabled (CPU may come soon)
- Learn about Deep Learning
 - Implement everything from scratch by ourselves

What we've achieved this week?

- Framework to create computation graphs
 - Simple feedforward NN
 - RNN, GRU, LSTM...
 - Binary or multiclass classifier
- Forward step
 - Classify, given input data and weights
- Backward step
 - Learn weights using backpropagation
- Tested with small datasets
 - MNIST (digit image recognition task)
 - MT
- Documentation
 - Doxygen

Compute graphs

What needs to be finished?

- Basic features:
 - Data shuffling
 - More random distributions
 - ...
- Model serialization & deserialization
- Documentation
- ...