096260 - Deep learning course:
Recurrent networks exercise

Submission date: 26/1/2017

1 Recurrent networks

Write a complete training procedure for a word-level LSTM recurrent network on Penn Treebank dataset. Design and train your network so that it will satisfy the 2 following goals:

- Final word-level perplexity on the test-set should be < 120
- Number of trainable parameters (weights) within the network should be $< 3M$

In addition to the requirements set, use the networks achieving these goals to generate 5 random sentences continuations to (no apostrophes): "Buy low, sell high is the" ...

1.1 Recurrent implementation

You may use any Torch recurrent implementation (or write your own).
Some popular choices:
Element-Research's rnn - Implementation shown in class
torch-rnn
eladhoffer recurrent.torch

1.2 Data

The Penn Treebank data (in text format, preprocessed to word-level) is available here under
ptb.train.txt
ptb.valid.txt
ptb.test.txt

You should train your network only with **ptb.train.txt**

1.3 Submission instructions

Submission will be in pairs (course partners) and will contain a short pdf report containing:

- Model architecture description, training procedure (data augmentation, regularization, optimization details etc).

- Two convergence graphs for your final model - for error and loss as a function of time (epochs). Each graph should depict both training and test performance.
- A short summary of your attempts and conclusions.

In addition, you should also supply:

- Code able to reproduce your results - we might test it on different variants on these datasets.

1.4 Grades policy

- Successful submission - 80 points.
- Report - 20 points.