Adam Goodge

Week 6: Vanishing Gradient Problem, Fancy RNN Reference material

Collection of the papers used for Week6: Vanishing Gradient Problem, Fancy RNN (GRU and LSTM) for reference:

1. 'On the difficult of training recurrent neural networks' (Pascanu et al., 2013) http://proceedings.mlr.press/v28/pascanu13.pdf

2. 'A simple way to initialize Recurrent networks of Rectified Linear Units' (Le et al., 2015) https://arxiv.org/pdf/1504.00941.pdf

3. 'Parsing with compositional vector grammars; (Socher et al., 2013) http://www.aclweb.org/anthology/P13-1045

4. 'Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation' (Cho et al., 2014) https://arxiv.org/pdf/1406.1078.pdf - GRU

5. 'An Empirical Exploration of Recurrent Network Architectures' (Jozefowicz et al., 2015) http://proceedings.mlr.press/v37/jozefowicz15.pdf - LSTM

6. 'LONG SHORT TERM MEMORY' (Hochreiter and Schmidhuber, 1997) http://www.bioinf.jku.at/publications/older/2604.pdf

7. 'Bidirectional Recurrent Neural Networks' (Schuster and Kuldip, 1997) https://pdfs.semanticscholar.org/4b80/89bc9b49f84de43acc2eb8900035f7d492b2.pdf