

Deep Learning for NLP

Recurrent Neural Networks



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Nils Reimers

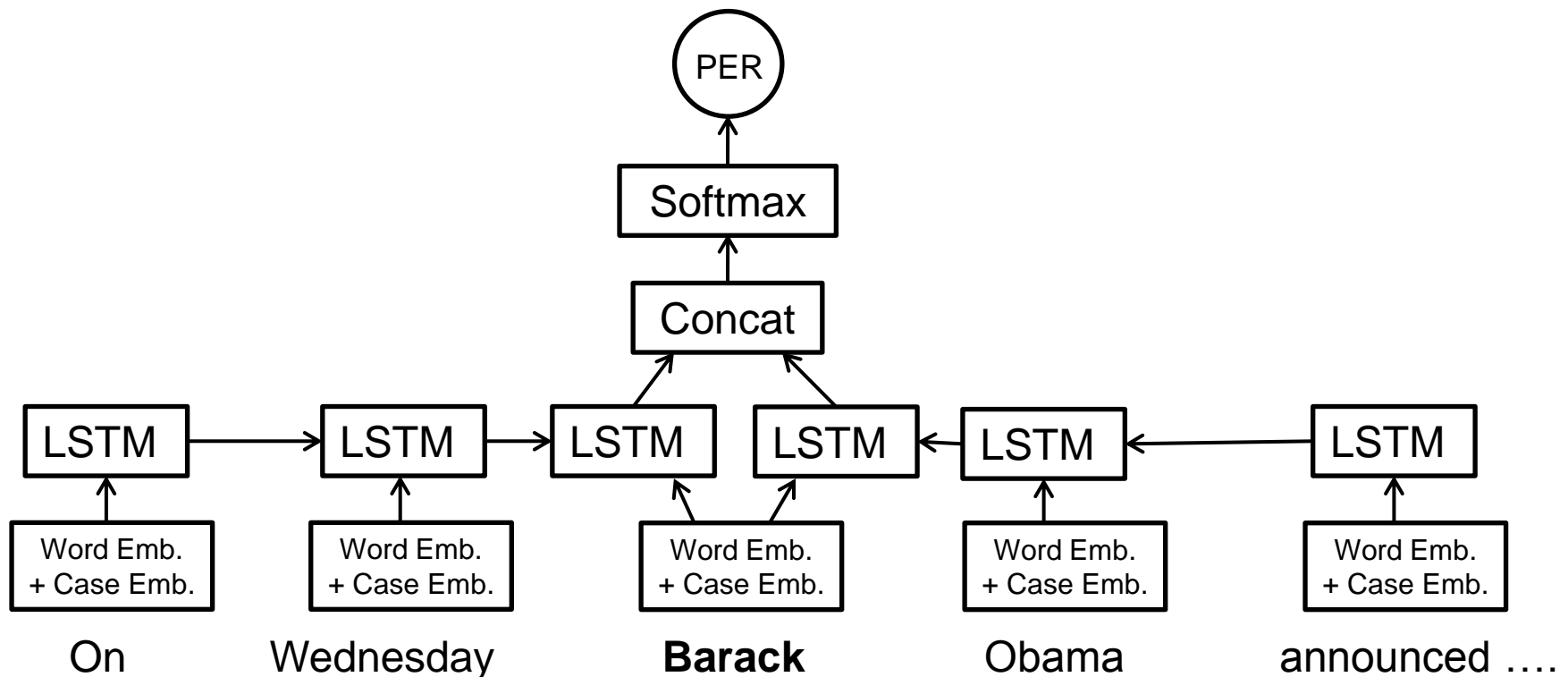
Course-Website: www.deeplearning4nlp.com

Recommended Readings

- This folder contains an implementation that is similar to Lample et al., 2016, *Neural Architectures for Named Entity Recognition*

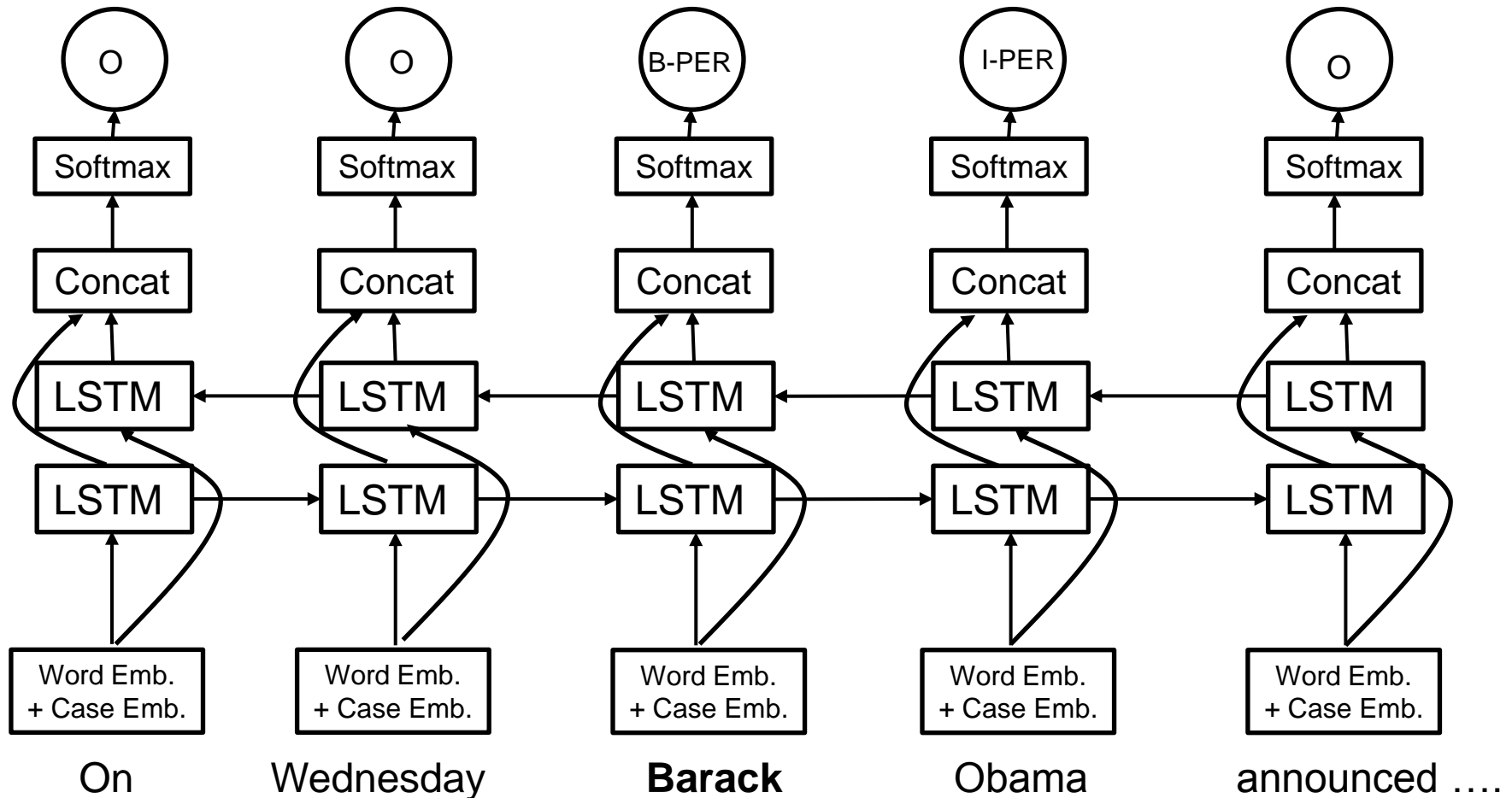
BiLSTM for Sequence Classification

- Forward LSTM reads all tokens until target token
- Backward LSTM reads all tokens until target token
- Output of both LSTMs are concatenated and fed into a softmax-classifier



BiLSTM for Sequence Classification

- We apply this structure to the whole sentence



- We apply online learning (training one sentence per time)
 - Through this trick, we don't need to pad our sentences
 - We must just Python lists and cannot use numpy arrays
- Word indices for the trainings data:
 - [[4, 5, 6], #First sentence, e.g. 'I am John'
 [5, 7, 9, 1, 2], #Second sentence, e.g. 'Hello PAUL MIKE JOHNSON !'
 ...]
- We converted the tokens to lower case, so we have case information:
 - [[1, 0, 1], #e.g.: initialUpper allLower initialUpper
 [1, 3, 3, 3, 4] #e.g.: initialUpper allUpper allUpper allUpper other
 ]
- Our labels look like:
 - [[0, 0, 1], #e.g.: O O B-PER
 [0, 1, 2, 2, 0], #e.g.: O B-PER I-PER I-PER O
 ...]