

Homework 09 – Language

Arthur J. Redfern
arthur.redfern@utdallas.edu

0 Outline

- 1 Reading
- 2 Theory
- 3 Practice

1 Reading

1. Language

Motivation: understand common uses of xNNs in language applications
https://github.com/arthurredfern/UT-Dallas-CS-6301-CNNs/blob/master/Lectures/xNNs_090_Language.pdf

Complete

2. Open sourcing BERT: state-of-the-art pre-training for natural language processing

Motivation: an alternative presentation of BERT
<https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>

Complete

3. The illustrated BERT, ELMo, and co. (how NLP cracked transfer learning)

Motivation: an alternative presentation of language modeling
<http://jalammar.github.io/illustrated-bert/>

Complete

2 Theory

None

3 Practice

4. Word embeddings. Understand the following examples and run them in Google Colab:

- https://pytorch.org/tutorials/beginner/nlp/word_embeddings_tutorial.html
- <https://gist.github.com/GavinXing/9954ea846072e115bb07d9758892382c>

Complete

5. Sentence embeddings (followed by additional tasks). Understand the following example and run it in Google Colab:

- https://pytorch.org/hub/huggingface_pytorch-transformers/

Complete

6. Language modeling. Understand the following example and run it in Google Colab:

- https://pytorch.org/tutorials/intermediate/char_rnn_generation_tutorial.html

Complete

7. Language translation. Understand the following examples and run them in Google Colab:

- https://pytorch.org/tutorials/beginner/torchtext_translation_tutorial.html
- https://pytorch.org/tutorials/intermediate/seq2seq_translation_tutorial.html

Complete

8. **[Optional]** Sequence to sequence models. A number of applications, especially in language processing, can be formulated as sequence to sequence tasks. For more information on this check out:

- <https://github.com/pytorch/fairseq>

Complete