

Nick

Week 7 - Advanced Attention

Week 7 - Advanced Attention

Slides

TOPICS

- Attention Code Example - Abhinav Kashyap @abhinav
- Attention Applications - Zining Zhang @zhangzn
- Scale up NMT- Nan Xiao @Nan
- Quasi-RNN - Adam Goodge @Adam Goodge

RESOURCES

- Pytorch tutorial of seq2seq with attention
- Google's Tensorflow tutorial of seq2seq with attention in the task of neural machine translation
- A nice interactive visualization of attention
- A paper that explains the role of the “grey box” from the sentinel network (slide 7)
- A Deep Reinforced Model for Abstractive Summarization
- Temporal Attention Model for Neural Machine Translation
- Pointer Sentinel Mixture Models On Using Very Large Target Vocabulary for Neural Machine Translation
- Get To The Point: Summarization with Pointer-Generator Networks
- BlackOut: Speeding up Recurrent Neural Network Language Models With Very Large Vocabularies
- Achieving Open Vocabulary Neural Machine Translation with Hybrid Word-Character Models
- Quasi-Recurrent Neural Networks