Nan

W5 Summary

LECTURE

Recurrent Neural Netowrks & Language Models

- Stanford Slides
- YouTube
- Presentation

TOPICS

- Language models - Kee Yuan Chuan @yuanchuan
- RNN language model - Jin Zhe @Jin Zhe
- Backprop - Wu Yizhuo @Yizhuo & Wu Jiacheng @Jiacheng
- RNN applications - Liu Juncheng @Juncheng

Q&A

- n-gram model probabilities approximation
    - independence assumptions are made so that each word depends only on the last n − 1 words
    - This Markov model is used as an approximation of the true underlying language
    - Smoothing is used for data sparseness problem
- softmax
    - Why do we use output values as exponent to the natural log base?
        - used to clip output values to ~ [0, 1]
        - we use exp so that we can push incomparable values close together for comparison (map +/- \infty into 0,1 similar to why we use sigmoid for other work)
    - non-negative value
    - + log(C) to avoid overflow of large number
        - Detailed explaination from CS231n

RESOURCES

- Understanding LSTM Networks
- Memory-Augmented Neural Networks
    - to memorize the context far away can be a hard staff. but here is a project on NLP doing the logic induction quite well
- Anyone Can Learn To Code an LSTM-RNN in Python
    - Simple implementation of RNN to predict addition problems.
- Lecture 10: Recurrent Neural Networks from CS231n
    - Many to One is described in slide 30
- The Unreasonable Effectiveness of Recurrent Neural Networks
- EINSTEIN SUMMATION IN DEEP LEARNING
    - More on the Einsum notation
- Kronecker delta
- Backpropagation with Softmax / Cross Entropy
- The Softmax function and its derivative
- Recurrent Neural Networks Tutorial, Part 3 – Backpropagation Through Time and Vanishing Gradients
- Styles of Truncated Backpropagation
    - interesting experiment on TBPTT
- Chenglei's Medium Posts
    - Backpropagation
    - Backpropagation Through Time

CODE EXAMPLES

- CharRNN (From Seedbank)
- Minimal character-level language model with a Vanilla Recurrent Neural Network (Gist from Karpathy)
- RNN example for forward & backward pass (Gist from Nan)
- Trump bot (Github Repo)