CS565: Intelligent Systems and Interfaces

Neural Language Modeling

Semester: Jan – May 2019

Ashish Anand

Associate Professor, Dept of CSE

IIT Guwahati

Announcements

- Scribe
  - Monika Khandelwal and Jayant Reddy: 20th Feb, 2019

Recap

- In the last lecture
  - Language Modeling: Definition
  - Language Modeling: N-gram Models
  - Language Modeling: Parameter Estimation

Objective

- Discuss Neural Language Model
  - What is generalization principle ?
  - Bengio et al. 2003 Model with softmax output layer of vocab size
  - Issue with Bengio et al. 2003 models
  - Hierarchical output layer and why it is advantageous.

References

- https://web.stanford.edu/~jurafsky/slp3/7.pdf
- Video Lectures of Hugo Leorchelle
  - 10.5 **https://tinyurl.com/y46lhrac**
  - 10.6 **https://tinyurl.com/yx958nmu**
  - 10.7 **https://tinyurl.com/y47oq4ck**

- Original Paper References
  - Bengio et al. A neural probabilistic language model JMLR 2003, [https://tinyurl.com/y62kna6u]
  - Morin and Bengio Hierarchical probabilistic neural network language model
  [https://tinyurl.com/y2avoeuz]