

EECS 496: SEMINAR IN STATISTICAL LANGUAGE MODELING

Winter 2018

Instructor:	Doug Downey	Email:	ddowney@eecs.northwestern.edu
Time:	Tuesday 1:00-3:30PM	Place:	Kresge Centennial Hall 2-435

Course Pages:

1. http://www.cs.northwestern.edu/~ddowney/courses/395_Winter2018/

Course Description: Statistical language models (SLMs) assign probabilities to sequences of words, and are used in systems that perform speech recognition, machine translation, and many other tasks. In recent years, deep neural networks have provided radically improved SLMs. This course will cover both the fundamental technologies that comprise statistical language models, from more classical n-gram models to recent memory-based deep neural networks, along with applications of the techniques to important tasks in artificial intelligence. Students will be required to read and present research papers, and to complete a substantial course project.

Office Hours: After class, or by appointment.

Course Materials: No textbook. The course material will be comprised of research papers in the field.

Course Goals: The goal of this course is to familiarize graduate students and advanced undergraduates with the current state-of-the-art in statistical language modeling. Students will read recently published papers in the field.

Prerequisites: EECS 349 or permission of the instructor.

Grading Policy: Leading a paper discussion (30%), Class participation (20%), Project (50%).

Course Objectives:

- Have a general understanding of the current state-of-the art in statistical language models.
- Understand how at least one statistical language model is implemented and can be applied (via the course project).
- Be able to understand, and think critically about, recent research papers in the field of statistical language modeling.

Background Resources:

1. [Deep Learning Textbook](#)
2. [Deep learning in neural networks: An overview](#) - (Schmidhuber; Science 2015)
3. [A Neural Probabilistic Language Model](#) - (Bengio et al; JMLR 2003)
4. [Advances in NLP](#) - (Hirschberg and Manning; Science 2015)
5. [Practical Recommendations for Gradient-Based Training of Deep Architectures](#) - (Bengio 2012)

Calendar:

- 1/16 Word Embeddings
[Linguistic Regularities in Continuous-Space Recurrent Neural Network Language Models](#) - (Mikolov et al., NAACL-HLT 2013)
[GloVe: Global Vectors for Word Representation](#) - (Pennington et al., EMNLP 2014)
Recommended: [Efficient estimation of word representations in vector space](#) - (Mikolov et al., 2014)
Supplement: [VecShare](#) - Fernandez et al., EMNLP 2017
- 1/23 Recurrent Neural Networks and LSTMs
[Recurrent neural network based language model](#) - (Mikolov et al; Interspeech 2010)
[Generating Sequences with RNNs](#) - (Graves 2013)
[Regularizing and Optimizing LSTM Language Models](#) - (Merity et al., 2017)
Supplement: [Resource to Understand LSTMs better; Long Short Term Memory](#) - (Hochreiter & Schmidhuber 1997)
- 1/30 Sequence to sequence, and attention
[Learning Phrase Representations using RNN EncoderDecoder for Statistical Machine Translation](#) - (Cho et al 2014)
[Sequence to Sequence Learning with Neural Networks](#) - (Sutskever et al NIPS 2014)
[Effective Approaches to Attention-based Neural Machine Translation](#) - (Luong, Pham, Manning 2015)
Supplement: [Resource to understand attention](#)
- 2/6 Image Captioning
[Show, Attend and Tell: Neural Image Caption Generation with Visual Attention](#) - (Xu et al; ICML 2015)
[Knowing When to Look: Adaptive Attention via A Visual Sentinel for Image Captioning](#) - (Metamind; 2016)
[Show and Tell: Lessons Learned from the 2015 MSCOCO Image Captioning Challenge](#) - (Vinyals et al., IEEE Trans. on Pattern Analysis and Machine Intelligence, 2017)
- 2/13 Conversation
[A Neural Network Approach to Context-Sensitive Generation of Conversational Responses](#) - (Sordoni et al., NAACL-HLT 2015)
[Generating High-Quality and Informative Conversation Responses with Sequence-to-Sequence Models](#) - (Shao et al., EMNLP 2017)
[A Knowledge-Grounded Neural Conversation Model](#) - (Ghazvininejad et al., AAAI 2018)

- 2/20 Memory Networks and QA
 - [End-to-end Memory Networks](#) - (Sukhbaatar et al., 2015)
 - [Key-value Memory Networks for Directly Reading Documents](#) - (Miller et al., EMNLP 2016)
 - [Question Answering on Knowledge Bases and Text using Universal Schema and Memory Networks](#) - (Das et al., ACL 2017)
- 2/27 Word Embeddings Revisited
 - [Improving Distributional Similarity with Lessons Learned from Word Embeddings](#) - (Levy et al., TACL 2015)
 - [Issues in evaluating semantic spaces using word analogies](#) (Linzen, Workshop on Evaluating Vector Space Representations for NLP, 2016)
 - [Evaluating vector-space models of analogy](#) - (Chen et al., CogSci 2017)
- 3/6 Text Generation Grab Bag
 - [Smart Reply: Automated Response Suggestion for Email](#) - (Kannan et al., KDD 2016)
 - [Learning to Generate Product Reviews from Attributes](#) - (Dong et al., EACL 2017)
 - [Context-aware Natural Language Generation with Recurrent Neural Networks](#) - (Tang et al., AAAI 2017)
 - [Definition Modeling: Learning to Define Word Embeddings in Natural Language](#) - (Noraset et al., AAAI 2017)
- 3/13 Project presentations