Course work plan

- Weekly lectures consist of theory lectures and practical lectures.
- Lecture slides are provided through Dongguk e-class site.
- Weekly schedule is as following (JK: Jihie Kim, KK: Kwangil Kim)

1. Word Embeddings - JK

Term Project Handout:

(1) Project Handout (Robust QA track)

(2) Project Handout (IID SQuAD track)

(3) End to End Question-Answering System Using NLP and SQuAD Dataset

Suggested Readings:

(1) Efficient Estimation of Word Representations in Vector Space (original word2vec paper)

(2) Distributed Representations of Words and Phrases and their Compositionality (negative sampling paper)

Additional Readings:

(1) GloVe: Global Vectors for Word Representation (original GloVe paper)

(2) Improving Distributional Similarity with Lessons Learned from Word Embeddings

(3) Evaluation methods for unsupervised word embeddings

(4) A Latent Variable Model Approach to PMI-based Word Embeddings

(5) Linear Algebraic Structure of Word Senses, with Applications to Polysemy

(6) On the Dimensionality of Word Embedding

2. Neural Networks - KK

Suggested Readings:

(1) CS231n notes on network architectures

(2) CS231n notes on backprop

Additional Readings:

(1) matrix calculus notes

(2) Review of differential calculus

(3) Derivatives, Backpropagation, and Vectorization

(4) Learning Representations by Backpropagating Errors (seminal Rumelhart et al. backpropagation paper)

(5) Yes you should understand backprop

(6) Natural Language Processing (Almost) from Scratch

3. Dependency Parsing - KK

Suggested Readings:

(1) Jurafsky & Martin Chapter 14

Additional Readings:

(1) Incrementality in Deterministic Dependency Parsing

(2) A Fast and Accurate Dependency Parser using Neural Networks

(3) Globally Normalized Transition-Based Neural Networks

(4) Universal Stanford Dependencies: A cross-linguistic typology

(5) Universal Dependencies website

(6) Dependency Parsing

4. Recurrent Neural Networks - JK

Suggested Readings:

(1) N-gram Language Models (textbook chapter)

(2) Sequence Modeling: Recurrent and Recursive Neural Nets (Sections 10.1 and 10.2)

Additional Readings:

(1) Sequence Modeling: Recurrent and Recursive Neural Nets (Sections 10.3, 10.5, 10.7-10.12)

(2) The Unreasonable Effectiveness of Recurrent Neural Networks (blog post overview)

(3) On Chomsky and the Two Cultures of Statistical Learning

(4) Learning long-term dependencies with gradient descent is difficult (one of the original vanishing gradient papers)

(5) On the difficulty of training Recurrent Neural Networks (proof of vanishing gradient problem)

(6) Vanishing Gradients Jupyter Notebook (demo for feedforward networks)

(7) Understanding LSTM Networks (blog post overview)

5. Seq2Seq Model and Neural Machine Translation - JK

Suggested Readings:

(1) Sequence to Sequence Learning with Neural Networks (original seq2seq NMT paper)

(2) BLEU (original paper)

Additional Readings:

(1) Statistical Machine Translation slides, CS224n 2015 (lectures 2/3/4)

(2) Statistical Machine Translation (book by Philipp Koehn)

(3) Sequence Transduction with Recurrent Neural Networks (early seq2seq speech recognition paper)

6. Attention Mechanism - JK

Suggested Readings:

(1) Neural Machine Translation by Jointly Learning to Align and Translate (original seq2seq+attention paper)

(2) Attention and Augmented Recurrent Neural Networks (blog post overview)

Additional Readings:

(1) Massive Exploration of Neural Machine Translation Architectures (practical advice for hyperparameter choices)

(2) Achieving Open Vocabulary Neural Machine Translation with Hybrid Word-Character Models

(3) Revisiting Character-Based Neural Machine Translation with Capacity and Compression

7. Transformer - JK

Suggested Readings:

(1) Attention Is All You Need

(2) Layer Normalization

Additional Readings:

(1) The Illustrated Transformer

(2) Transformer (Google AI blog post)

(3) Image Transformer

(4) Music Transformer: Generating music with long-term structure

8. Pretrained Language Models - JK

Suggested Readings:

(1) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

(2) Contextual Word Representations: A Contextual Introduction

Additional Readings:

(1) The Illustrated BERT, ELMo, and co.

(2) Martin & Jurafsky Chapter on Transfer Learning

9. Mid-term Exam - JK & KK

10. Question Answering - JK

Suggested Readings:

(1) Dense Passage Retrieval for Open-Domain Question Answering

(2) Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks

Additional Readings:

(1) SQuAD: 100,000+ Questions for Machine Comprehension of Text

(2) Latent Retrieval for Weakly Supervised Open Domain Question Answering

(3) Learning Dense Representations of Phrases at Scale

(4) Bidirectional Attention Flow for Machine Comprehension

(5) Reading Wikipedia to Answer Open-Domain Questions

11. Representing and Using Knowledge in NLP - JK

Suggested readings:

(1) ERNIE: Enhanced Language Representation with Informative Entities

(2) Language Models as Knowledge Bases?

Additional Readings:

(1) Barack's Wife Hillary: Using Knowledge Graphs for Fact-Aware Language Modeling

(2) Pretrained Encyclopedia: Weakly Supervised Knowledge-Pretrained Language Model

12. Chatbots and Dialog System - KK

TBD

13. Natural Language Generation – KK

Suggested readings:

(1) The Curious Case of Neural Text Degeneration

(2) How NOT To Evaluate Your Dialogue System

Additional Readings:

(1) Get To The Point: Summarization with Pointer-Generator Networks

(2) Hierarchical Neural Story Generation

14. Information Extraction - JK

TBD

15. Project Report & Evaluation - JK & KK