CS 6956: Deep Learning for NLP Fall 2019

Homework 1

Handed out: 4 February, 2019
Due date: 20 February, 2019

This homework asks you to read two papers and comment upon them.

1. In class, we saw that `word2vec` employed the idea of negative sampling to estimate its parameters. In this question, you will go into the details of this technique. In particular, you will explore the connection between negative sampling and a different estimation technique called *noise contrastive estimation*.

   This connection is explained in following technical note:

   > Dyer, Chris. "Notes on noise contrastive estimation and negative sampling." arXiv preprint arXiv:1410.8251 (2014).[1]

   (a) [2 points] Explain why techniques such as noise contrastive estimation or negative sampling are needed for training word embeddings? Is the problem still relevant given modern GPU based architectures?

   (b) [3 points] In section 3, the note mentions that negative sampling can be seen as a special case of noise contrastive estimation if certain conditions are met. Show that this argument is correct by working out the math.

2. [5 points] In class, we saw that word embeddings are dense, low-dimensional embeddings of text. An alternative approach could be to use what are sometimes called "traditional NLP features", which are sparse and high dimensional. A natural question could be about how these two methods compare with each other.

   Read the following paper:

   > Levy, Omer, and Yoav Goldberg. "Linguistic regularities in sparse and explicit word representations." In Proceedings of the eighteenth conference on computational natural language learning, pp. 171-180. 2014.[2]

   Recall the discussion in the class about how to review papers. The following questions will guide you through the process.

---

[1]Available online: `https://arxiv.org/abs/1410.8251`
[2]Available online: `http://www.aclweb.org/anthology/W14-1618`

(a) Identify the main scientific contribution of the paper

    i. A good paper will try to answer a specific question. Try to identify this question. For example, this could be a new idea, a unification of previous ideas into a more general framework, or an implementation of an idea to show that it works.

    ii. Sometimes the question asked by the paper might be its most important contribution

    iii. Sometimes the novelty of the approach could be its biggest selling factor

Write a brief summary of the contribution(s) of the paper in your own words. This should be a short paragraph (about 3-4 sentences).

(b) Identify the positive aspects of the paper. What are the good ideas in the paper? What is their potential influence? How could you generalize the ideas? If you start research in this direction, what will be your next step?

(c) Does the paper have any weaknesses? Is there something fundamentally incorrect in the paper? Suppose you had to change something in the paper that made the contributions more clear. What would the changes be? Different experiments? More experiments? More theory? Other changes?

(d) Does the paper use the right set of experiments? Is the experimental setup convincing? What other experiments/results would strengthen the paper? Do the experiments really persuade you about the papers claims? Do the results persuade you?

(e) End with an open-ended discussion about the novelty, validity and (potential) impact of the paper.