

Week 5

11-777

Due: Thursday, Sept 28, by 4:30 pm

Instructions

A small yet significant part of the course is answering a few questions based on a paper you just read. By design, some of these questions are open-ended and the purpose is to start a thinking process. Some of those questions may already have been answered in recent literature (or in the given paper itself). We encourage you to, both, come up with your own ideas and survey recent papers before answering these questions.

A few key points to take care of:

- Answer **two out of the three questions** (you are welcome to answer all 3), unless stated otherwise.
- A lot of questions ask you to suggest changes or make alterations. It would be nice to support it with logical / mathematical arguments. Cite all sources used in the process of coming up with your answer. Figures and equations, if they support your arguments will be appreciated.

Questions

Combining Language and Vision with a Multimodal Skipgram Model

1. What are the drawbacks of MSDMs highlighted in the paper? Does the MMSkip-Gram model address these? How?
2. No direct visual evidence was used for some words in the dataset under consideration in the paper. These words are thus influenced by vision only by propagation. How do you think this indirect propagation occurs? Is it effective?
3. What are the differences in results obtained for concrete and abstract terms as described in the paper? Can you think of a better way to inject visual or other semantically meaningful evidence for abstract terms like 'wrong' and 'right'? How can you extend the given model to incorporate another modality (say speech)?