Identifying our Group’s Research Problem:

Part 1:

The high-level problems:

Gender-bias in Natural Language Processing problem

How do we debias gender bias in word embeddings?

Specific technical problems:

In GloVe and Word2Vec, how does the gender bias form? Where do they come from? How do we train the word embeddings and debias them?

What is the relationship between word embeddings and coreference resolution?

Part 2:

ERSP problems:

Gender-bias in Natural Language Processing problem

How do we debias gender bias in word embeddings?

Where does the gender bias come from?

Part 3:

I really enjoy the process of learning and discussing. We find papers about gender bias in NLP problems, read them and bring problems to discuss during the meetings. I have learned a lot by looking up for videos and documents about the concepts and math.

Yet what I like the least is that I have to figure out the math in the papers to actually understand the debiasing algorithms. I don’t know how in depth should we know about the math to complete the tasks in the following months.

I’m confused about what we are going to do in the following weeks. Are we going to figure out a debiasing algorithm on our own?