Sai Teja Paladi

1. Read paper - [10 + 10 + 10 = 30]

a) According to you, why does this behavior of word embeddings happen when running analogies?

I think the unintended gender stereotypes that were found in embeddings are because gender is present in society and gender-based biases affect large groups of people. I think this is the major reason we see gender stereotype along with gender-appropriate analogies.

b) What approach is proposed in the paper to mitigate it?

A methodology for modifying an embedding to remove gender stereotypes while maintaining desired associations is proposed. Defined two kinds of bias, direct and indirect bias. The debiasing is done by neutralizing where the components are removed along gender subspace and equalize, i.e., any gender-neutral word is equidistant to the pairs of gender-specific words.

 c) Do you think the approaches actually mitigate? Any problems you anticipate in practice?

Yes, the approach of debiasing mitigates the gender bias in embeddings, and the embeddings can be used in the applications.

2.Create and run notebook [10 + 10 + 20 + 20 = 60]
 Create your own copy python notebook and execute from the tutorial python notebook at:
https://github.com/PLN-FaMAF/Bias-in-wordembeddings/blob/main/main_tutorial_bias_word_embedding.ipynb
 Activities to do are:
a. load word embedding
b. do visualization
c. run analogies: examples and your own (at least 3)
d. run one mitigation method

Link to your completed notebook is at:
https://github.com/BunnyTeja/NLP/tree/main/Quiz-4/Code

3.Apply to your project [3 + 4 + 3 = 10]
 Now consider your course project.

a. Does gender bias in word embedding relevant to your project? How?

 Yes, it is relevant to my project because I am generating query-based text summarization, and it should return the answer without any gender bias.

b. If yes, what strategy will you use to improve the situation?

 I can use the approach mentioned above to de-bias the word embeddings by neutralizing and equalizing. So that I can preserve the useful properties and reduce the gender bias in embeddings.

c. If no, what ethical issue(s) do you anticipate for your project and the strategy one may use?