Advanced NLP

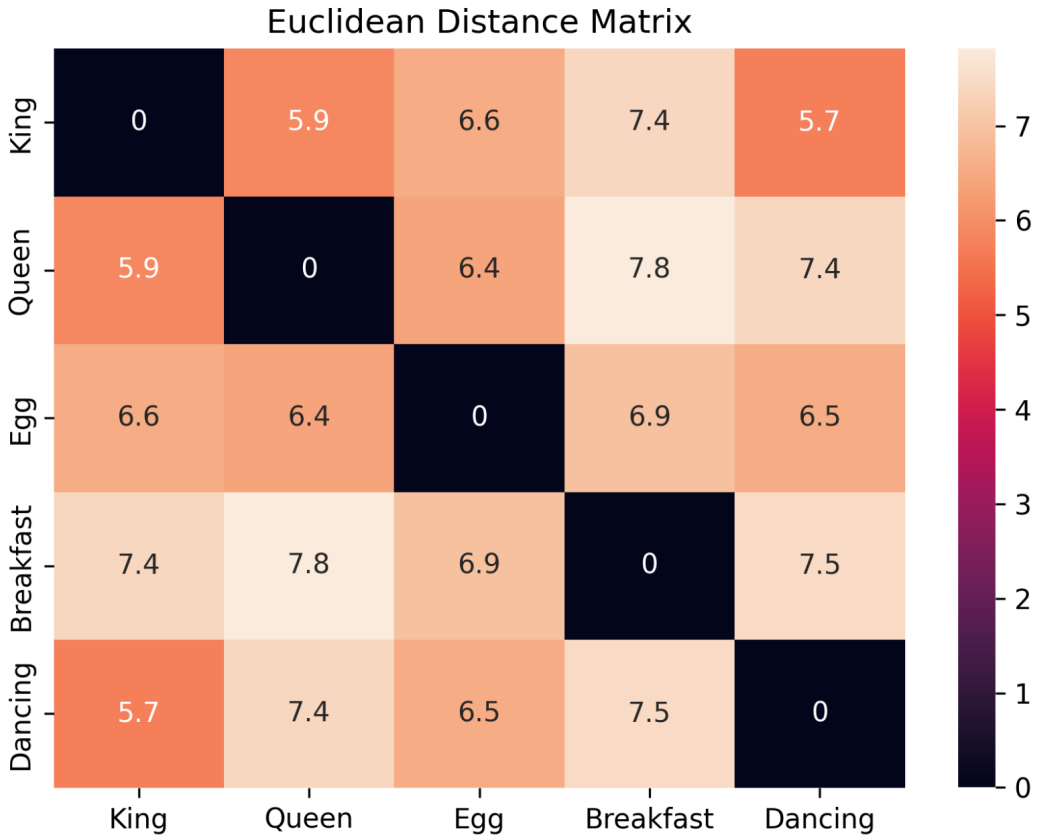
Assignment 2 - ELMo

T. H. Arjun, 2019111012, CSD

All the assignment requirements were implemented including:

- **Implementation of the ELMo Language Model as described by the Paper with a character level embedding with BiLM with Source Code.**
- **The model was implemented from scratch in Pytorch and trained on the given corpus.**
- **More Details in README.md**

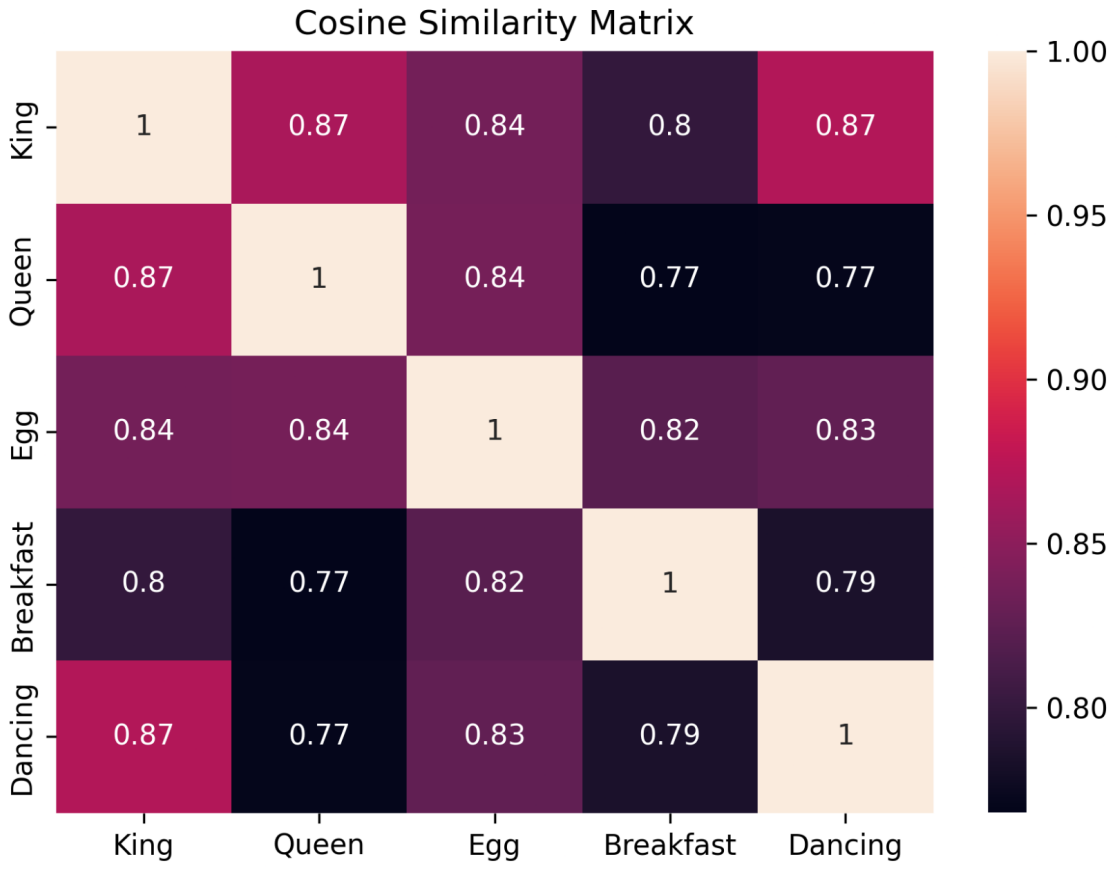
1. Calculate Euclidean distance for any 5 different pairs of words.
   Below I have plotted the heatmap matrix of Euclidean Distance between 5 words from the ELMo Trained as part of the assignment.



Notable Observations as expected:

- King is very near to Queen
- Breakfast is very near to Egg
- Dancing is very near to King and Queen

2. Calculate cosine distance for any 5 different pairs of words.
   Below I have plotted the heatmap matrix of Cosine Similarity between 5 words from the ELMo Trained as part of the assignment.



Notable Observations as expected:

- King is very similar to Queen
- Breakfast is very similar to Egg
- Dancing is very near to King and Queen

Contextual Word Embeddings

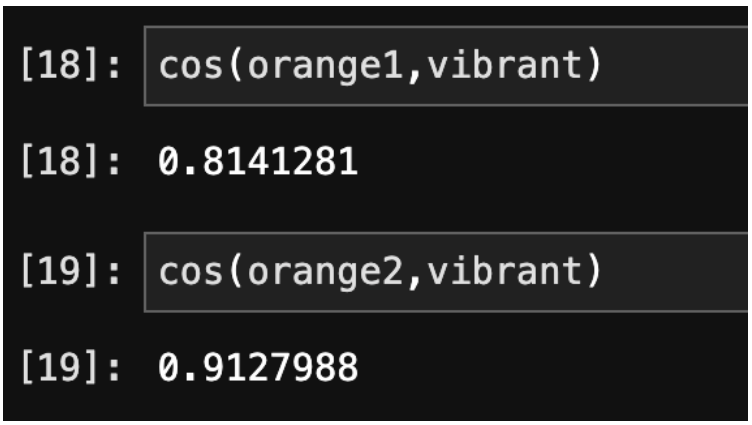
It should be considered that the above vectors represented are contextual and are dependent on the sentence. This is the advantage of ELMo as it produces contextual word embeddings.

Now let us consider two sentences:

- “The orange from the supermarket was juicy and tasty. Their fruits are the best.”
- “The artist painted his house with a tint of orange color and it looked vibrant”

The word orange has different meanings in both these sentences where, in the former, it’s referring to fruit while in the latter it is referring to a color. Hence the representation should be different. Now let us consider the embeddings of these two oranges with the embedding of the word “vibrant” in the context “The painting has a lot of vibrant colors and it looked beautiful.”

-



  We see that the cosine similarity of the word embedding of orange in the context of colour is more similar to the word vibrant than when it is in the context of fruit as expected.