Hiva Mohammadzadeh

Report for HW 2

HW2a:

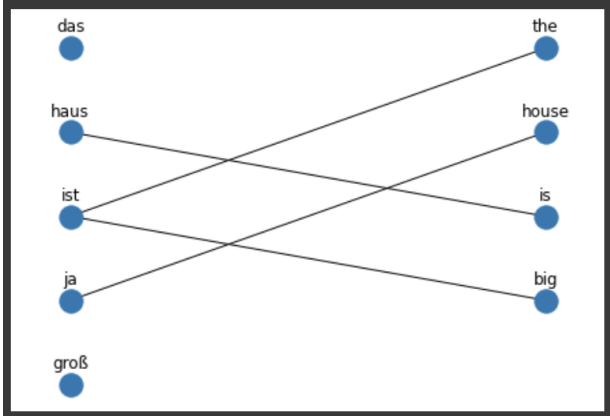
The IBM Alignment model calculates the probability of a source word mapping to all the words in the target based on the given training data. It calculates the transnational probabilities and basically just takes the target word that has the highest probability to be the translation and answer. The alignment model is a statistical model, so it just looks calculates the probabilities of each word in the input and maps each word to it's corresponding translation accordingly.

Some reasons why it does well sometimes is because the data was good and out of the given data, those source words were perfectly aligned with their corresponding target word. This is because in the training data, a source word always showed up to map to the same target word. It could also be that the same exact sentence that we're trying to translate showed up in the training set. On the other hand, some reasons why it doesn't do well for the sentence that I chose is because there are multiple source words mapping to the same target words. Also, it could be that the context that the word had showed up in the training set was different than the context that we're looking for.

For Example:

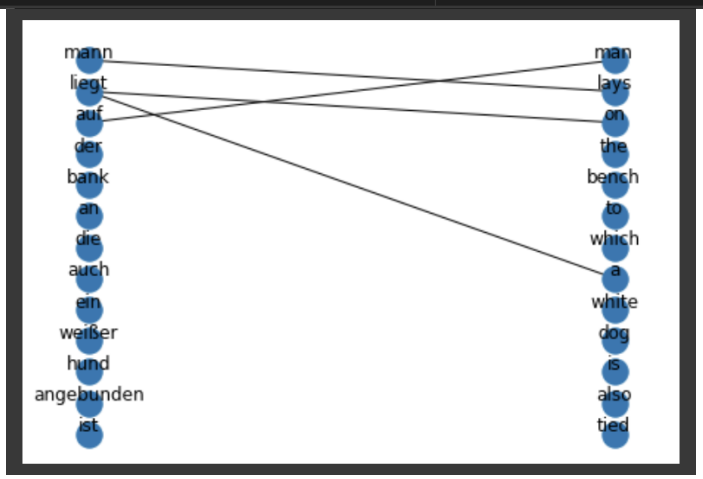
This is an example where the model does good. This is right according to a German dictionary. For a sentence like this where the model has been trained on most of the words before and has seen the words in the same context before, it is really easy for it to make the right prediction.

```
visualize_alignment(alignment, ['das', 'haus', 'ist', 'ja', 'groß'], ['the', 'house', 'is', 'big'])
```



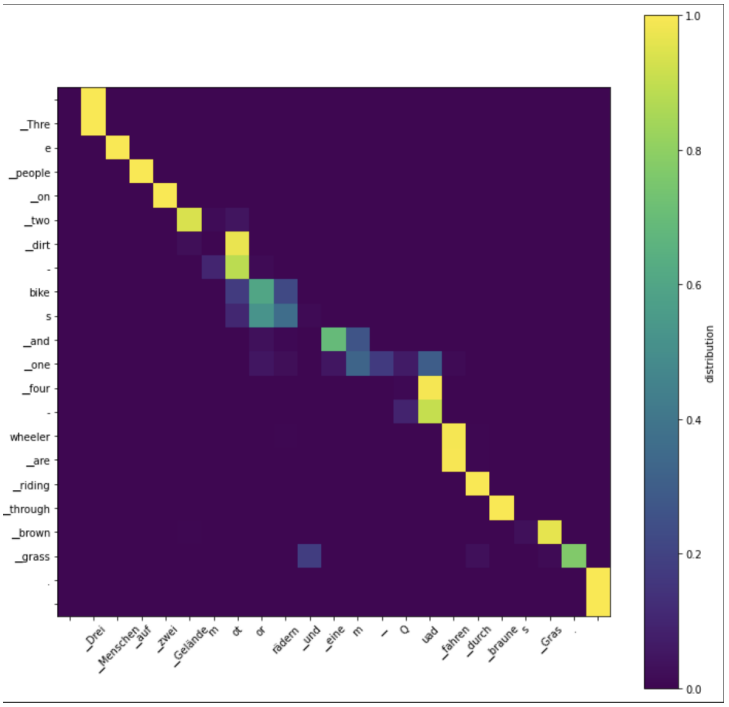
On the other hand, this is an example where this alignment model fails. As we can see, it did not even classify most of the words that are in the sentence. This is because it has never seen that word, so it cannot align it to an extent where it's reasonable.

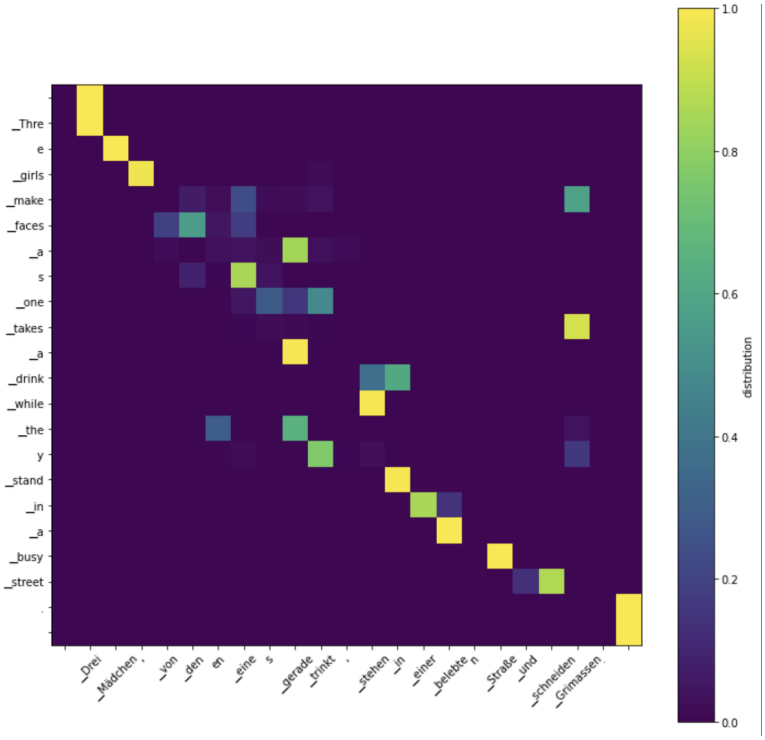
```
visualize_alignment(alignment,['ein', 'mann', 'liegt', 'auf', 'der', 'bank', 'an', 'die', 'auch', 'ein', 'weißer', 'hund', 'angebunden', 'ist'],
                    ['a', 'man', 'lays', 'on', 'the', 'bench', 'to', 'which', 'a', 'white', 'dog', 'is', 'also', 'tied'])
```

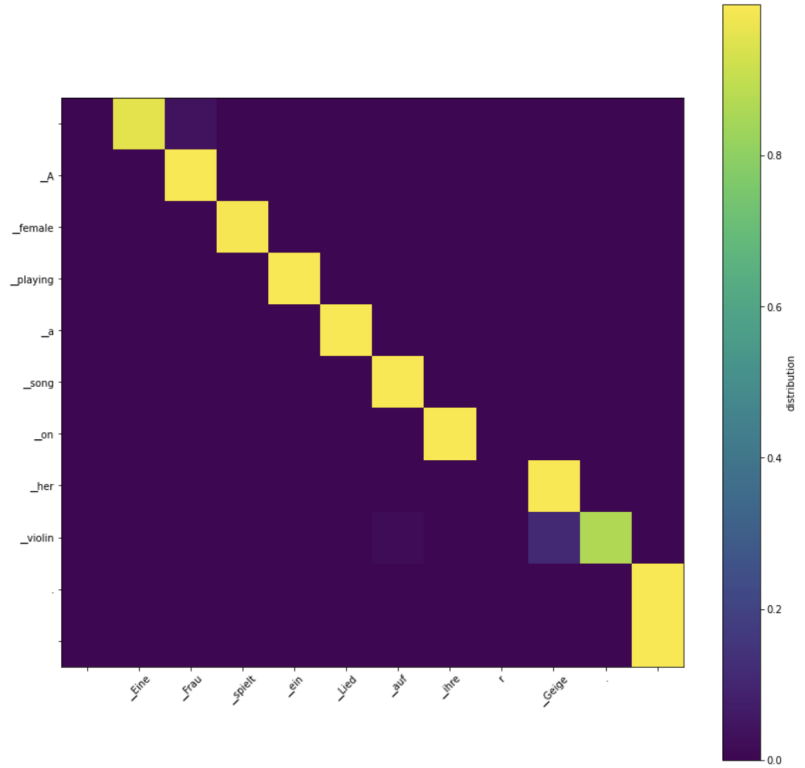


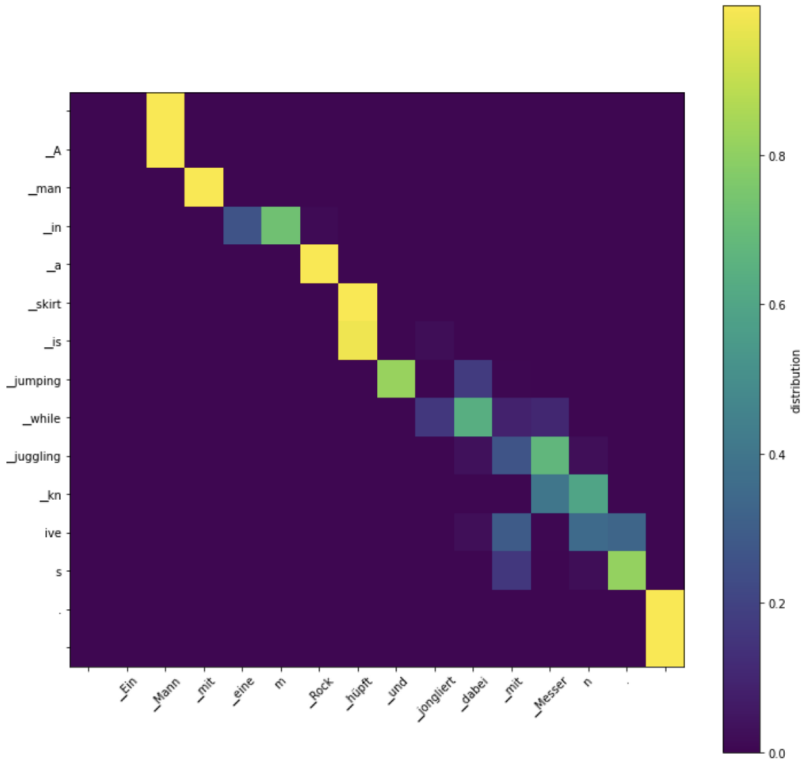
HW2b: Report: generate attention visualizations for at least four sentences and describe general trends you observe in your report. What similarities or dissimilarities do you see between the attention maps and the alignment visualizations from Part A?

I looked at a lot graphs in order to accurately estimate the behaviour and characteristics. In these plots we can see that attention distribution shows us a much better visualization of what's going on with our predictions. We can also see that even though our model is pretty good and we can see that a lot of the times that the intensity of the probability distribution is perfect, for some words that don't make sense to be by themselves (like "and" and "the") in a sentence without any context, it has a bad distribution. It perfectly classifies the words that are uniques like numbers and nouns.






_A
_female
_playing
_a
_song
_on
_her
_violin
.
_Eine
_Frau
_spielt
_ein
_Lied
_auf
_ihre
r
_Geige
.
0.0
0.2
0.4
0.6
0.8
distribution


_A
_man
_in
_a
_skirt
_is
_jumping
_while
_juggling
_kn
ive
s
.
_Ein
_Mann
_mit
_eine
m
_Rock
_hüpft
_und
_jongliert
_dabei
_mit
_Messer
n
.
0.0
0.2
0.4
0.6
0.8
distribution