

Assignment 3

Hasan Mustafa
AIL-721

November 27, 2024

1 Overview

All code was run on CPU(either HPC or local). Stable baseline was run on Google Colab.

1.1 Quick Navigation Guide:

Trained a single layer autoregressive decoder transformer with

- Single Attention Layer
- Masked Attention
- Residual Connection
- Position Encoding

2 Data Set

Amazon Reviews Dataset was used. Specifically 500 rows of the full text column of the dataset were used which was about 200K tokens with a vocabulary size of 5K tokens.

3 Embeddings

We use 50 dimensional Glove embeddings from Glove Twitter, it has originally 1M tokens, we filter out only the tokens in our vocabulary. The embeddings were made trainable.

4 Summary Statistic

We plot perplexity vs training iterations. Final perplexity for model with 4 heads is 4.153233. Final perplexity for model with 12 heads 3.790837.

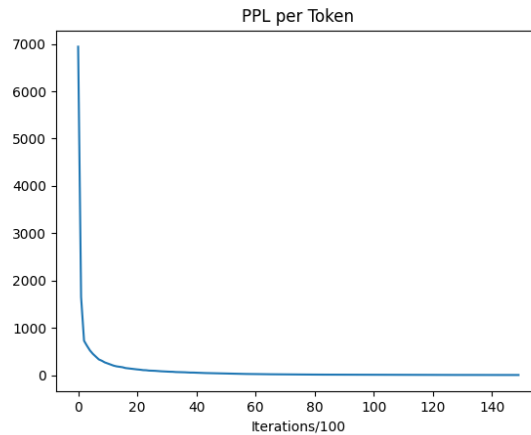


Figure 1: PPL vs iterations

5 Visualization

Implemented a mechanism to access attention map in the attention heads. Plotted graph between top 15 strongest connection in the attention map, additionally we plot the attention scores as a heatmap.

6 Example Sentences

Example sentence for 4 head model.

"it i went on line to amazon and not only found it but found it in the canister and cheaper than buying all the individual packets we like this better because you can scoop and measure easier i would highly recommend this product just save some for me **cant have any time we like this new in this particular about half the baklava i used**"

"i would consider this a true mens product and would recommend it without reservation i had read that monin grenadine contained no artificial coloring or preservatives as stated in the product description and other websites had stated the same thing not true at all i just received my bottle in **the mail and read the honey store could want soothing pallet cleansing quality of aloe vera** "

Example sentences for 12 head model.

in half for smaller dogs those are all the reasons my dog likes these treatsbr br the reason i like them organic ingredients make these treats healthy for my dog i love that i can give her a healthy treat that she lovesgreat tasting organic product i have been **more organic diet i was almost too a 10 year custody battle needless to say it**

they shipped me a case overnight this mustard is so good that i need to find things to put it on a nice hot pretzel would be perfecttry these we love them if you have a costco close by you may find them cheaper if not this is your **time to then if you use half my wife drinks 45 cups i do not need**

=

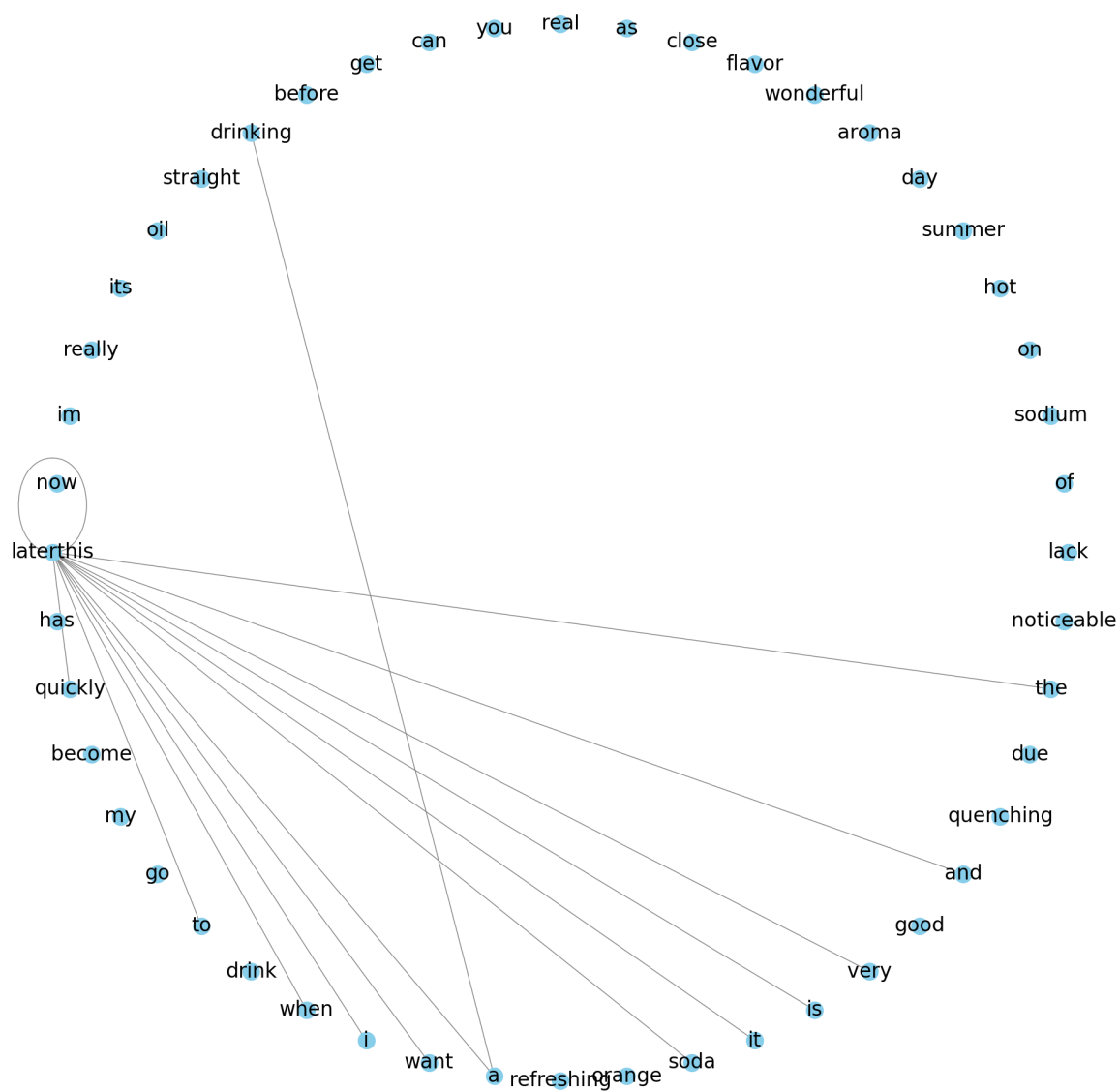


Figure 2: Graph from Attention Map

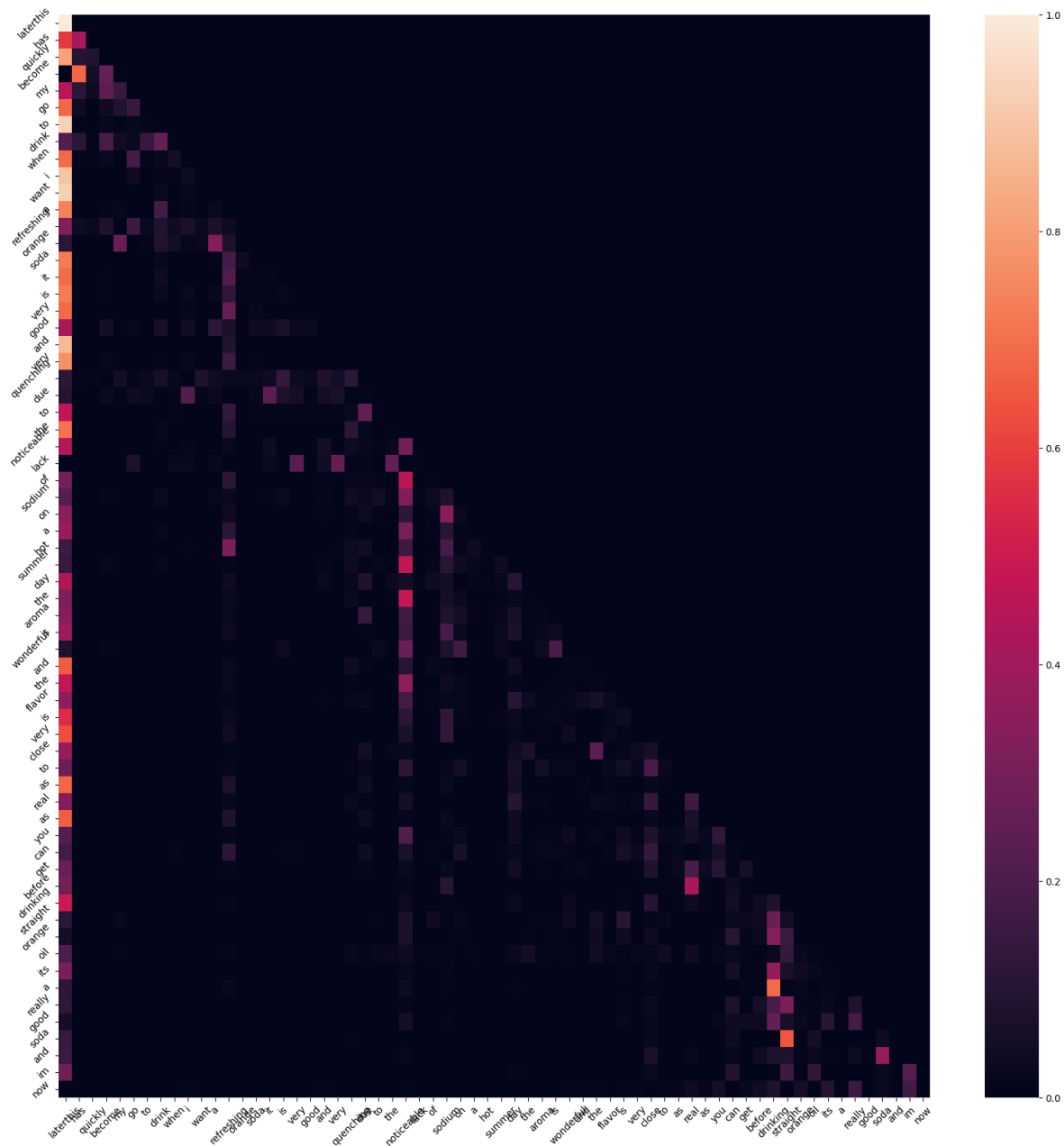


Figure 3: Heatmap Attention