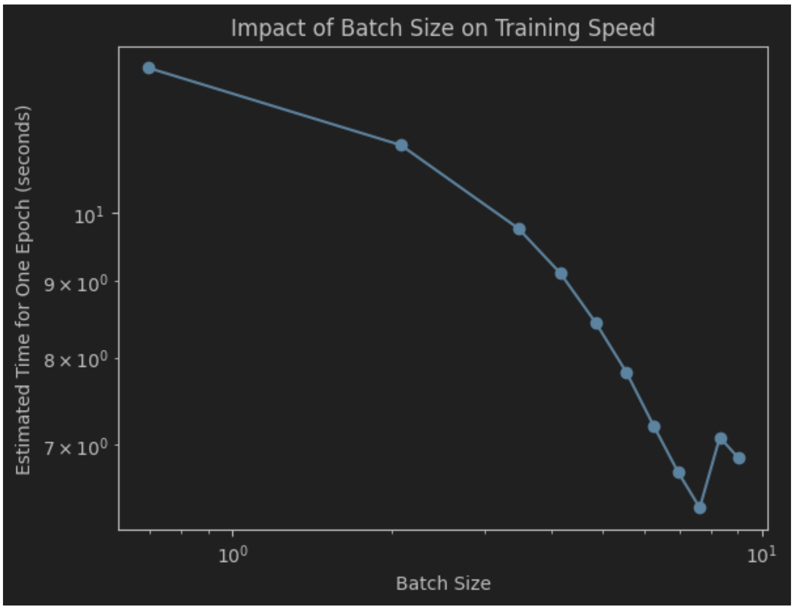
Problem 12:

Plot of batch size:

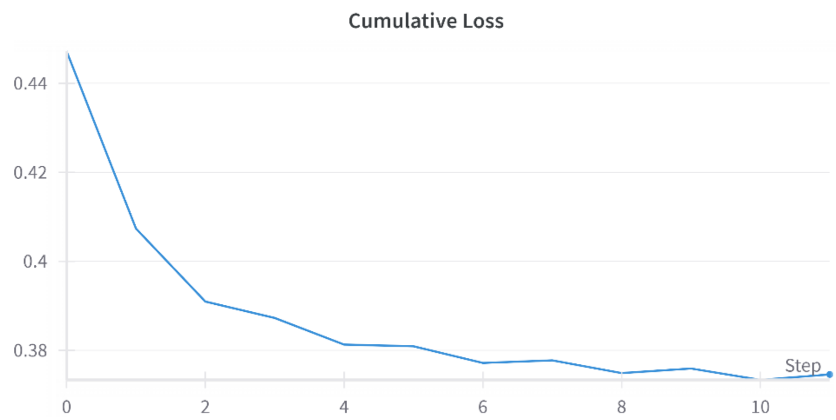


Description of what I see: Before batch size reaches 2048, as the batch size doubles, the estimated time will approximately become half. However, after that the time will not become less and fluctuate within a small range. I think the reason might be that 4096 is larger than my computer's max handling ability at the same time.

What batch size would I choose: 2048

Problem 13:

Plot of wandb:



Problem 15:

Results:

Most similar words to 'books': [('novels', 0.772312581539154), ('ione', 0.7335072159767151), ('trigiani', 0.7068877220153809), ('rosato', 0.6997941136360168), ('trilogys', 0.6966578960418701), ('thors', 0.6965705156326294), ('trilogies', 0.6898884177207947), ('patricelli', 0.6856436729431152), ('sandra', 0.6826285719871521), ('fossum', 0.6789857149124146)]

Most similar words to 'photos': [('photographs', 0.8947330713272095), ('pictures', 0.8914361000061035), ('pics', 0.8500222563743591), ('drawings', 0.8424339890480042), ('reproductions', 0.8058515191078186), ('illustrations', 0.7983523011207581), ('designs', 0.7664406895637512), ('paintings', 0.7623581290245056), ('photo', 0.744396984577179), ('stitching', 0.7426446080207825)]

Most similar words to 'man': [('woman', 0.9297835230827332), ('lady', 0.764117419719696), ('girl', 0.747891366481781), ('boy', 0.7254368662834167), ('widow', 0.7097799181938171), ('biologically', 0.7035444974899292), ('warrior', 0.6889328956604004), ('guy', 0.6875222325325012), ('men', 0.6839041709899902), ('murderous', 0.6816152930259705)]

Most similar words to 'football': [('baseball', 0.7944216728210449), ('hockey', 0.720067024230957), ('soccer', 0.7171085476875305), ('bachelor', 0.696029543876648), ('basketball', 0.6940032243728638), ('commando', 0.6905590891838074), ('tennis', 0.6902896761894226), ('sports', 0.6650619506835938), ('filmmaker', 0.6444071531295776), ('csi', 0.6343899965286255)]

Most similar words to 'son': [('daughter', 0.9354985952377319), ('grandson', 0.881630003452301), ('granddaughter', 0.8378354907035828), ('niece', 0.8368213176727295), ('yr', 0.7918955683708191), ('neice', 0.7663732767105103), ('sister', 0.7555080652236938), ('nephew', 0.7518795728683472), ('youngest', 0.7482140064239502), ('mom', 0.7328506708145142)]

Most similar words to 'fantasy': [('sci', 0.883333146572113), ('fi', 0.8693403601646423), ('dystopian', 0.8524616956710815), ('genre', 0.8329605460166931), ('scifi', 0.818652331829071), ('romance', 0.8167240619659424), ('horror', 0.8118466734886169), ('steampunk', 0.8045616149902344), ('fiction', 0.7922078371047974), ('paranormal', 0.7827293872833252)]

Most similar words to 'magnificent': [('stunning', 0.7854254841804504), ('geoff', 0.7655529379844666), ('incredible', 0.7541369199752808), ('superb', 0.751526415348053), ('beautiful', 0.7426850199699402), ('amazing', 0.7088311314582825), ('fantastic', 0.7027348279953003), ('marvelous', 0.6956594586372375), ('outstanding', 0.6938199996948242), ('masterful', 0.692801296710968)]

Most similar words to 'active': [('compliance', 0.6461794376373291), ('participation', 0.642973005771637), ('ensuring', 0.6320324540138245), ('surge', 0.6155261993408203), ('tcm', 0.6027130484580994), ('entering', 0.5971044898033142), ('aviator', 0.5960518717765808), ('computerized', 0.5945338010787964), ('gymnast', 0.5927485823631287), ('navy', 0.5896925926208496)]

Most similar words to 'insight': [('insights', 0.8578913807868958), ('insite', 0.7659927606582642), ('glimpse', 0.7376072406768799), ('incite', 0.7328647971153259), ('penetrating', 0.6904246211051941), ('perspective', 0.687146008014679), ('understanding', 0.6617640852928162), ('wisdom', 0.6392198801040649), ('guidence', 0.6340915560722351), ('background', 0.6208731532096863)]

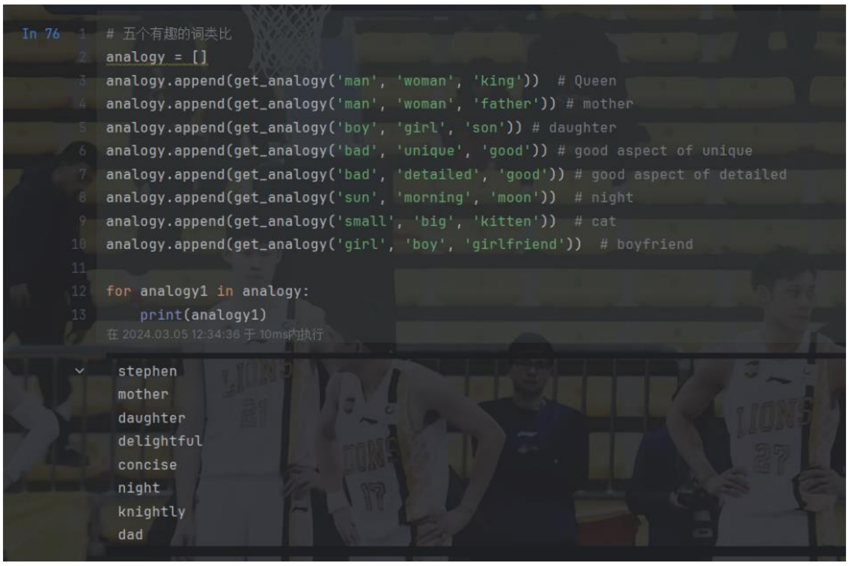
Most similar words to 'super': [('very', 0.6358646154403687), ('pretty', 0.6333441734313965), ('extremely', 0.6208550333976746), ('ridiculously', 0.5822411775588989), ('eater', 0.5794768333435059), ('incredibly', 0.5789140462875366), ('duper', 0.5772424936294556), ('alphas', 0.5744730234146118), ('unbelievably', 0.5731501579284668), ('insanely', 0.5690271258354187)]

Description of my findings:

The results highlight both the strengths and limitations of the model. While it successfully identifies semantically related words for common terms, showing a good grasp of their meanings and relationships (e.g., 'books' with literary terms, 'man' with human-related terms), it struggles with less frequent words. This may indicate the model's difficulty in finding similarities for words with fewer occurrences in the data, reflecting a limitation in capturing nuances for rarer terms. The findings suggest a balance between effective semantic understanding for common words and challenges in accurately representing less frequent terms.

Problem 16:

Results:



```
In 76
# 五个有趣的词类比
analogy = []
analogy.append(get_analogy('man', 'woman', 'king'))  # Queen
analogy.append(get_analogy('man', 'woman', 'father')) # mother
analogy.append(get_analogy('boy', 'girl', 'son')) # daughter
analogy.append(get_analogy('bad', 'unique', 'good')) # good aspect of unique
analogy.append(get_analogy('bad', 'detailed', 'good')) # good aspect of detailed
analogy.append(get_analogy('sun', 'morning', 'moon'))  # night
analogy.append(get_analogy('small', 'big', 'kitten'))  # cat
analogy.append(get_analogy('girl', 'boy', 'girlfriend'))  # boyfriend

for analogy1 in analogy:
    print(analogy1)
在 2024.03.05 12:34:36 于 10ms内执行

stephen
mother
daughter
delightful
concise
night
knightly
dad
```

```
In 79
analogy = []
analogy.append(get_analogy('hot', 'cold', 'summer'))  # Winter
analogy.append(get_analogy('seed', 'tree', 'egg'))  # Chicken or Bird, depending on the analogy intended
analogy.append(get_analogy('water', 'ice', 'steam'))  # Liquid, considering water's states of matter
analogy.append(get_analogy('cat', 'kitten', 'dog'))  # Puppy
analogy.append(get_analogy('write', 'writer', 'paint'))  # Painter
analogy.append(get_analogy('speak', 'speaker', 'teach'))  # Teacher
analogy.append(get_analogy('run', 'runner', 'swim'))  # Swimmer
for analogy1 in analogy:
    print(analogy1)
在 2024.03.05 12:38:41 于 9ms内执行

tuscany
siamese
hitch
threesome
artist
teaching
diver
```

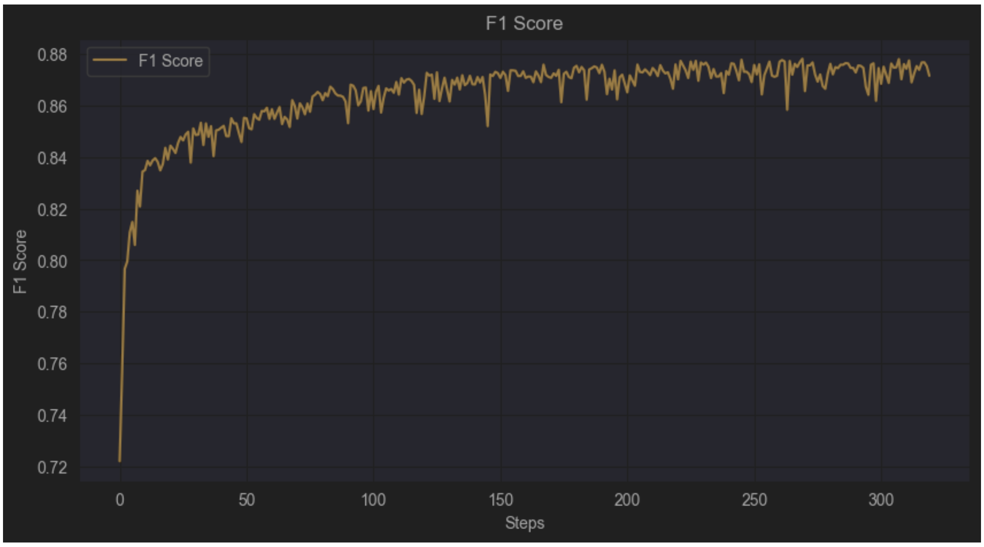
Description of my findings:

When dealing with the simple binary problem of things, the model's results are not bad. However, if we want to use it in more complexed problems, the model is too weak (maybe it's because I have used large to train it, not huge).

Problem 18:

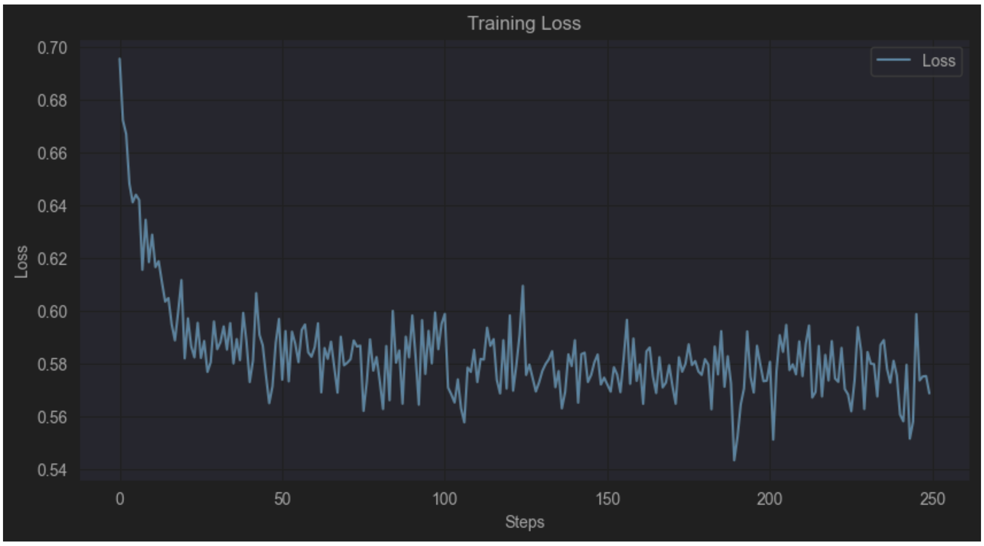
Plot of loss:

Plot of F1:

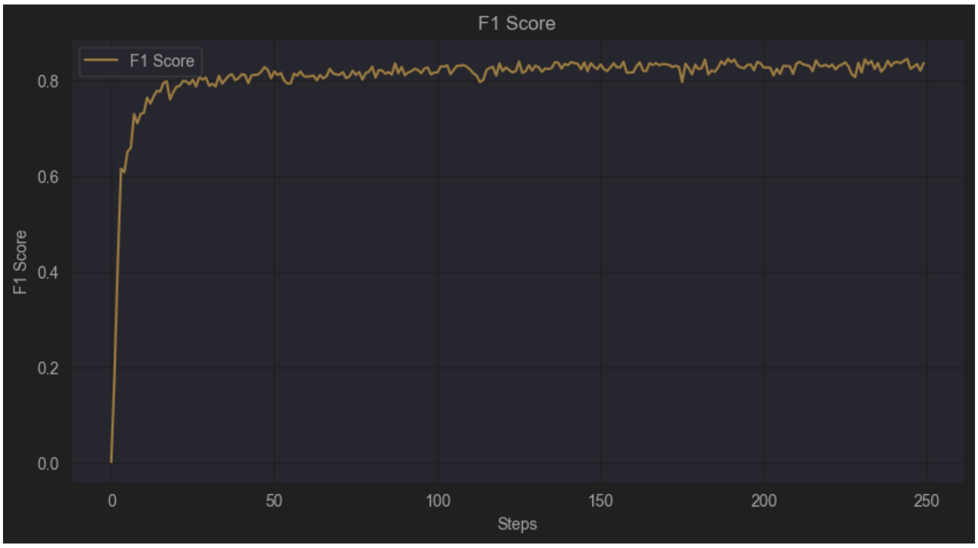


Problem 19:

Plot of loss:



Plot of performance:



Describe whether I think we should freeze our word vectors in this setting or not: Based on the results, the running time of freezing word vectors is much less than that of not freezing (freezing 20min compared to not freezing 50min), while the difference of final performance of model, i.e. F1score, is very small (freezing 0.836 compared to not freezing 0.876). Therefore, I think if we just focus on the accuracy of prediction and don't care about the running time, then we should not freeze the word vectors. However, since training takes a very long time if we don't freeze, I still suggest that we should freeze the word vectors when we are not at the final modification of the code. (Freeze to see if the code works and not freeze when submitting the final version of code)

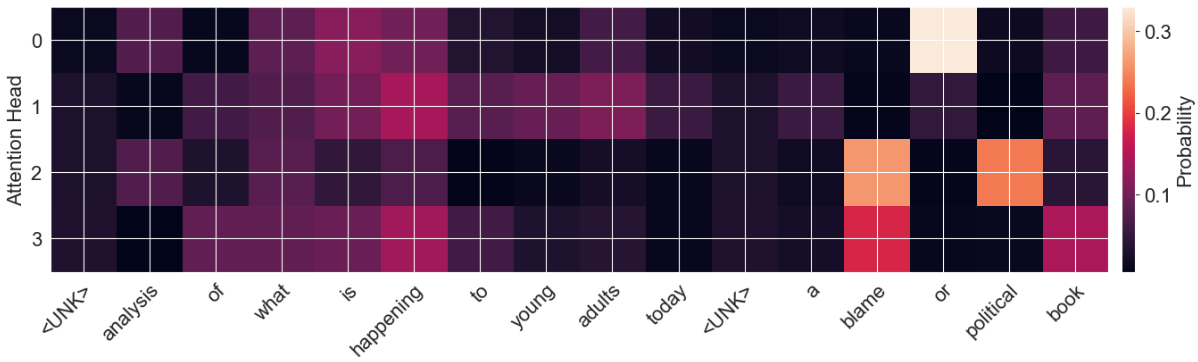
Problem 20:

Kaggle user name: Yanzhuo2001

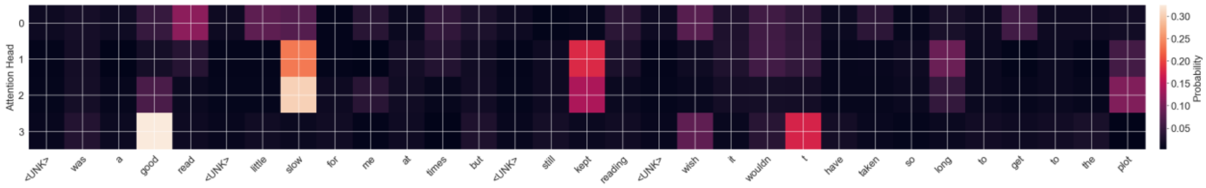
Problem 21.1:

Four attention plots:

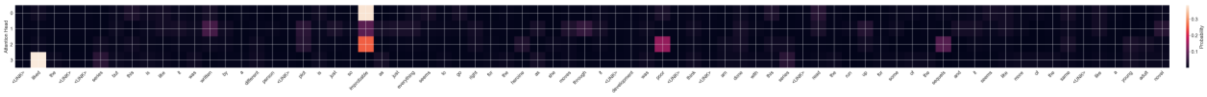
1. *Predicted Label: 1.0, Actual Label: 1*



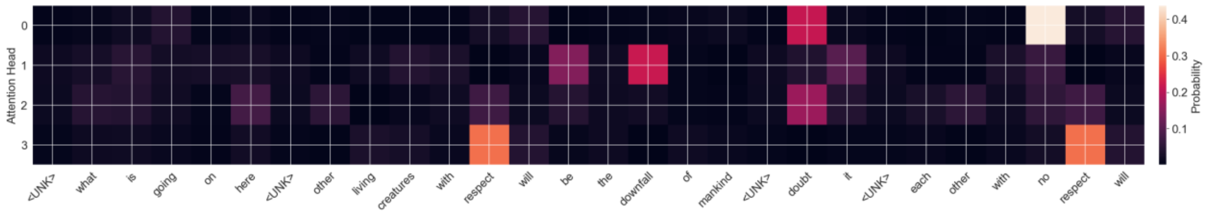
2. *Predicted Label: 1.0, Actual Label: 0*



3. *Predicted Label: 1.0, Actual Label: 0*



4. *Predicted Label: 1.0, Actual Label: 0*



Description for why I think the plots are interesting: I have chosen 2 plots with correct prediction and 2 plots with wrong prediction. And I found that many predictions pay much attention on stop words like "no" in 4, "or" in 2; while other predictions pay much attention on sentiment words like "liked" in 3, "good" in 1.

Problem 21.2:

Shot paragraph for what the attention heads are looking for: They are trying to focus on different kinds of words like stop words, sentiment words, subject and object words, etc.

Description for any differences I see between heads and whether there are any patterns in terms of what they focus on: In this example, I found that my head0 mostly focus on the stop words as figure2 and figure 4 show, while my head3 mostly focus on the sentiment words as figure 1 and figure 3 show. As for head1 and head2, I can't summarize the their patterns even if I look 10 more plots as shown in my .ipynb file.

Problem 21.3:

Description for whether the attention is good explanation and looking at the right thing: I think basically speaking the result is good since the model have paid attention to word "funny", which is an important sentiment word. However, the reason for wrong result might be that the model failed to pay attention to word "inspirational", which is also an important sentiment word.