UNIVERSITY OF ATHENS
DEPARTMENT OF INFORMATICS AND TELECOMMUNICATIONS

Deep Learning for NLP

Student name: <Αγγελος Κοντος>
sdi: <sdi2000089>

Course: *Artificial Intelligence II (M138, M226, M262, M325)*
Semester: *Fall Semester 2023*

Contents

1. Abstract

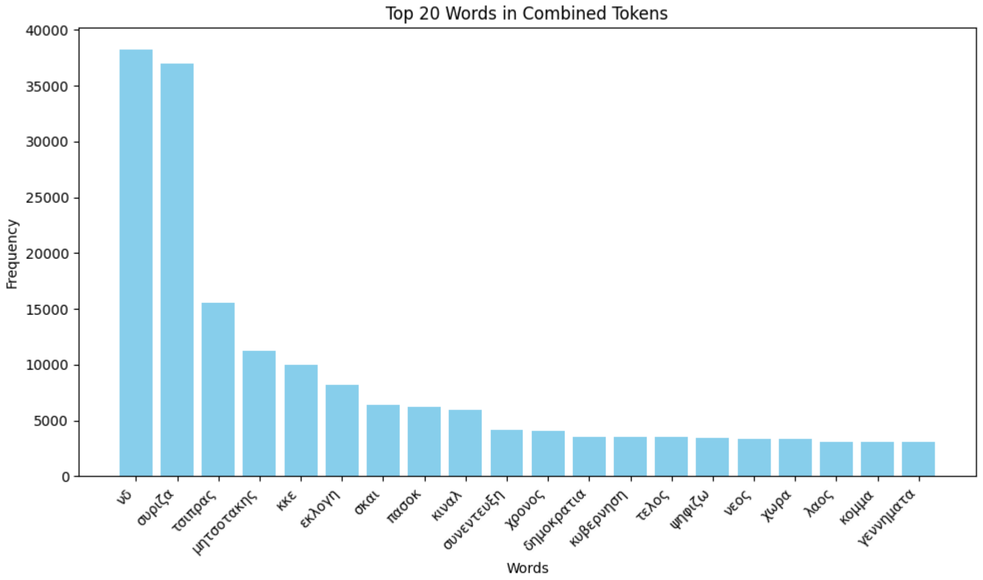
Essentially, the problem we are doing is the same as the poems we have to take a series of data that are tweets where they have positive, neutral or negative comments to flatten a neural system and get a result on new data, the difference is that now we have been asked to put memory (rnn, lstm, gru) in the neural system and attention, more or less I started playing which memory mode to choose lstm, rnn etc. it didn't have much variety (none) I also played with the layers which also didn't have variety and then I put attention .

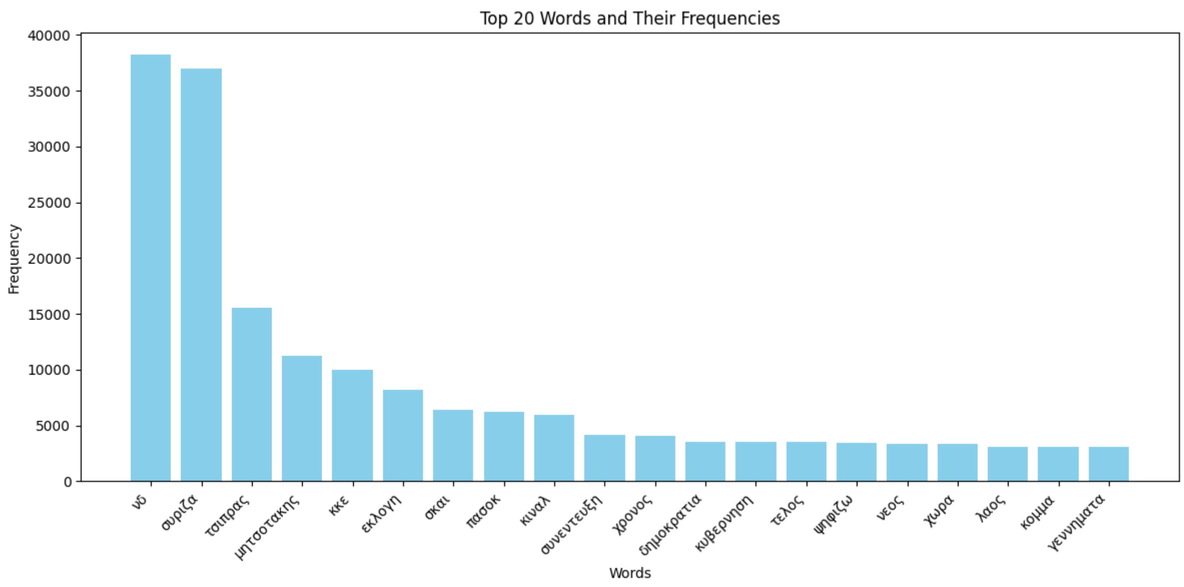
2. Data processing and analysis

2.1. Pre-processing

in the pre process I put lemetazation, I took out links, English letters, tones, I also joined as much as I could the words of parties, syriza tsipras, etc., I also took out #, @ and other symbols

2.2. Analysis





2.3. Data partitioning for train, test and validation

because I use optuna and dataloader every time part of the total material goes through, but I saw that the best is between batch$_{s}ize40 - 70orso$

2.4. Vectorization

I use word2vec where the exercise from before gave 100 size

3. Algorithms and Experiments

3.1. Experiments

Specifically, the experiments I did are from lstm 3 layers to 2 layers to 1 layer then I tried after seeing that the layers are not affected I put rnn and gru if something will change since it didn't change I played a little with the text editor to see if this is the problem, it didn't have anything, then I increased the batch$_{s}izeandreducedtheepochs, Ididn'tseemuchd$

1)3 *levels h and one LSTM*
2)3 *levels h and one LSTM*
3)3 *levels h and one LSTM*
4)2 *levelh and one LSTM*
5)2 *levels h and one LSTM with big batch_size and smaller epoch and all the*
6)1 *level h and one LSTM*
7)1 *level h and one RNN*
8)1 *level h and one GRU*
9)1 *level h and one GRU and new text editor*
10)1 *level h and one GRU and new text editor*
11)2 *level h and* 2 *LSTM with bigger h*
12)1 *level h attention and lstm*
13)1 *level h attention and lstm and diffrent text editor*
(*here Iput the union words the comatose*)

14)1 *level h attention and lstm and diffrent text editor and no party's (here Iput the conjunction words the politicalpartys)*

15) same as 14) just more in optuna experiments

Trial	Negative	Neutral	Positive	Score
1	46%	0%	30%	34%
2	45%	0%	41%	36%
3	41%	44%	0%	36%
4	41%	0%	45%	36%
5	41%	0%	45%	36%
6	40%	0%	46%	36%
7	41%	0%	45%	36%
8	40%	0%	46%	36%
9	42%	0%	44%	36%
10	42%	0%	45%	36%
11	30%	8%	44%	36%
12	46%	0%	39%	36%
13	49%	30%	0%	36%
14	30%	36%	0%	33%
15	45%	0%	44%	36%

Table 1: Trials

3.1.1. Table of trials.

3.2. Hyper-parameter tuning

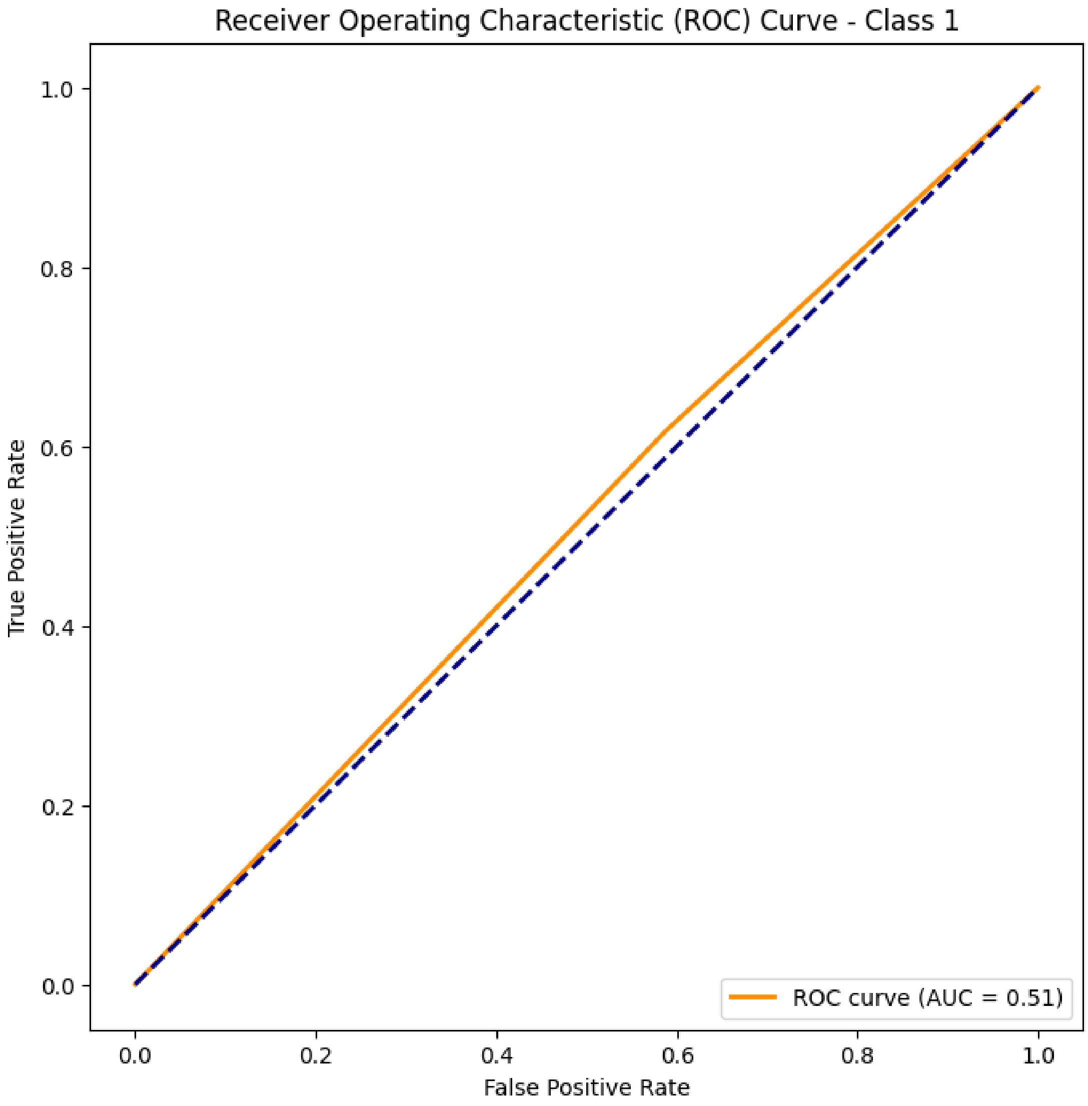
I would say it is under fitting because it is never at fault for all three qualities, either positive-negative or positive-neutral or neutral-negative, this from what I can see is due to the dataloader but then even extra effort to make sure that it will have all of its categories objects again this produced, and also the not good dataset

3.3. Optimization techniques

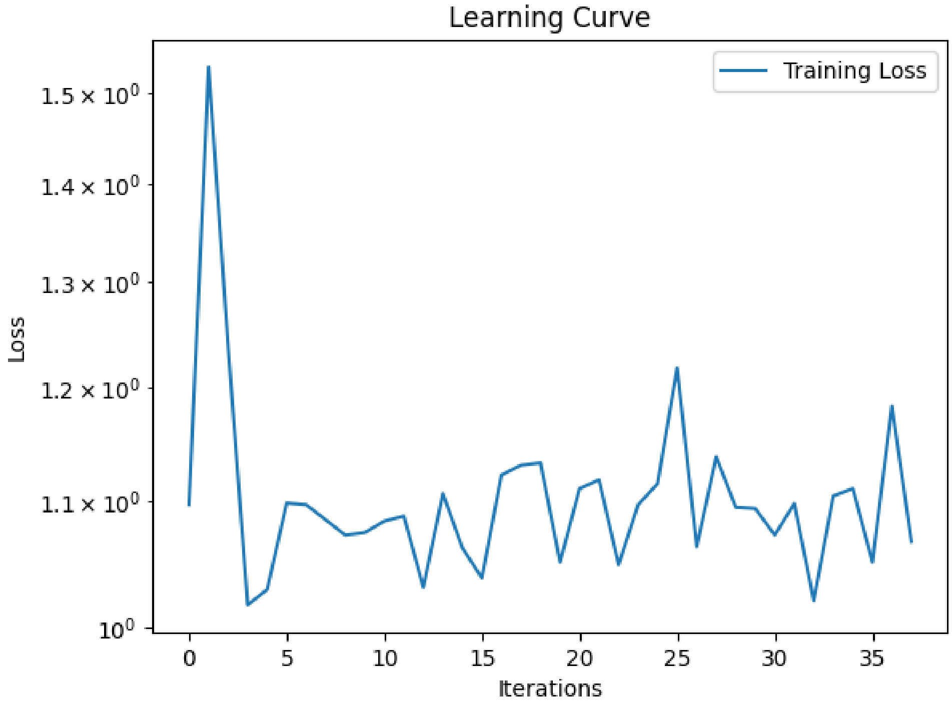
I used optuna changing first how much to reduce each layer how much lr, batch_size, epoch sir, i also changed twice the text editor maybe i could somehow make it better nothing happened, and the way i convert word2vec from sum to mean term, but general nothing was helping it

3.4. Evaluation

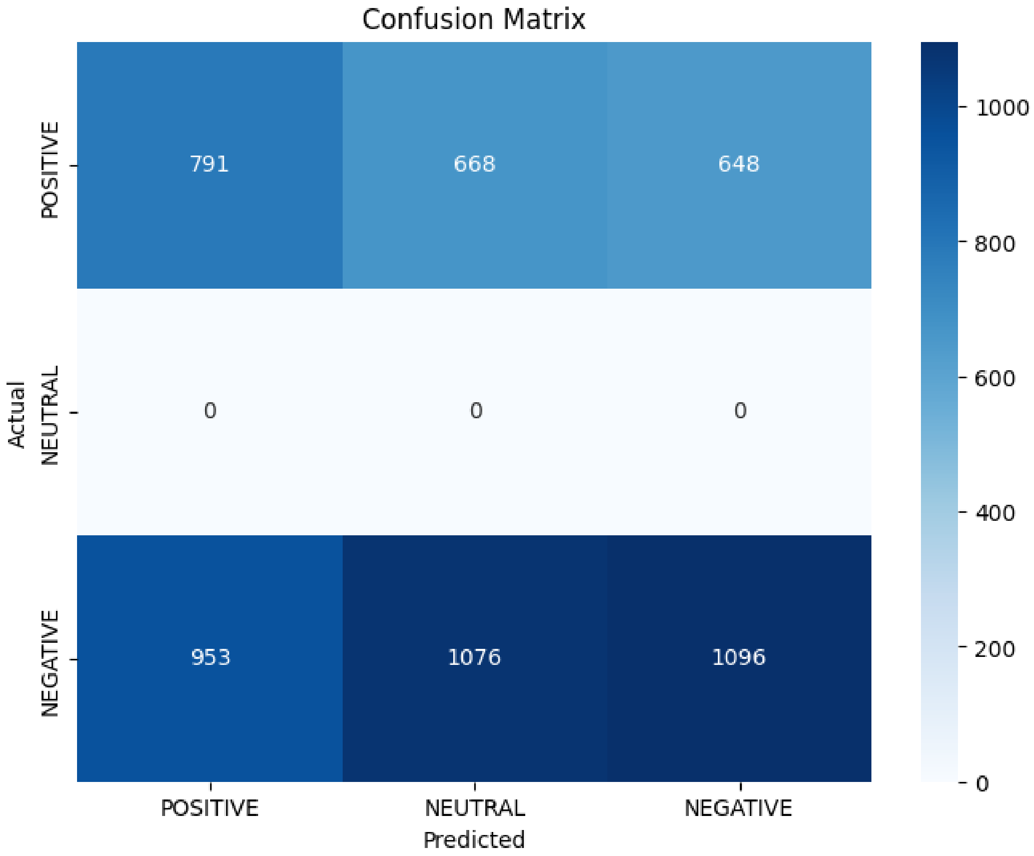
3.4.1. ROC curve. xa



3.4.2. Learning Curve. xa



3.4.3. Confusion matrix. xa



	POSITIVE	NEUTRAL	NEGATIVE
POSITIVE	791	668	648
NEUTRAL	0	0	0
NEGATIVE	953	1076	1096

4. Results and Overall Analysis

4.1. Results Analysis

are quite negative results shorter than the resulting tasks (and I don't know why I think it's the text editor but I haven't found a way to get results but changes to it), the only experiments I'd like to do more are with the text editor but also that I have asked other fellow students for ideas they have not done different things lemmetazion, links, English, tones etc etc

Trial	αρνητικα	ουδετερα	θετικα	Score
15	42%	0%	45%	36%

Table 2: Trials

4.1.1. Best trial.

4.2. Comparison with the first project

I found for some reason so far the best results are the first task I suspect because I mentioned it above but I don't know what it could be (it will be an incredibly small detail I believe)

4.3. Comparison with the second project

it is quite similar to the second task, it just wanted conversions to be able to do the rnn and the attention part, of course the results for me are a little worse, unfortunately the reason is that I probably wrote it before something will not have been done correctly or well in the text editor

5. Bibliography

I didn't use any extra material except the basics of pytorch for how the commands are made, plus the slides of the course, I didn't include everything but the most basic ones that I looked at.

<attention pytorch> <rnn pytorch> <example pytorch>