

CS 6957: NLP with Neural Networks Fall 2023

Mini Project 3

1. Number of parameters
- = no. of unique 4-grams + no. of unique 4-grams
- = 793900 + 149579
- = 943479.
- Perplexity on test set = 17.852444360206473
2. Training LSTM with weighted losses:
- a. One-layer LSTM
- i. Test Perplexity: 1.0682
- ii. Number of parameters: 279986
- b. Two-layer LSTM
- i. Test Perplexity: 1.5962
- ii. Number of parameters: 601586
3. The 4-gram model has thrice as many parameters as the 1-layer LSTMs model and one and a half times as many as 2-layer LSTM’s. But still we’re seeing such a jump in the test perplexity of LSTMs from the 4-gram model. I think this is because LSTMs see much larger context than the 4-gram model. But something interesting to notice here is that 2-layer LSTM is performing worse than the 1-layer LSTM! Although it is worth noting that, these LSTM models were trained using teacher forcing and the performances were also measured using teacher forcing. So, the real performance might be way worse than what is reported here and the answer of question 4 demonstrates that.
4. Predicting next 200 characters using sampling strategy:
- a. 1-layer LSTM
- i. The little boy was a hone opener under . eathor anone out the stard that the see the adicain on the recease to let to the eneter in the entine day a restal the add or as where roet to the deer til caten stranted the ha
- ii. Once upon a time in the thut an an note to strents or on the sitters on is with a tende in anidertan inders me not of the word or that lade air a creaty and with and the tentle earthand lire dees on the stadiied the enti
- iii. With the target in the teed he sing otherelboce not a hat first in the see a continp not the derited that the other the doen set is of three litted to the anting the his essine and the little recouting a other the roid
- iv. Capitals are big cities. For example, that treated to ketter the entieute on there three said a testerl to ere one in a home tuan grant muse in our great peasined with his master of the svensing of the as to be an that in the semotiness
- v. A cheap alternative to with hear the tine of the mastep be a men of that seen en, in the should to bare, and the benda no see law to an the and cradical hered wron a shrie an ard on the sea in Chipil with the two entle a t
- b. 2-layer LSTM
- i. The little boy wass se n rtts htutan h n k a eao er a o a i sni e nrls ira u ee oe e e a n t set co te os a nen t nhlei e d gne hst s mi e pe ta a e oeera en nr oe a as n e
- ii. Once upon a time in a i teir c soaacetf eicheu t e nee srll t epnsr etalnt toiete etii t ian er eo r s ,edr eer . t e e iaa a ste n o rai is tr r a di nse i e e eec eso to e eat et
- iii. With the target inr e r theeo e n r i eo e aeo i e h e tielo a rn e e lo naas ri iueea nage e e o eea i,nse th e a tntro nr u u nate a ns ltte ae td e sian ath i n e a tad n ne ma ea
- iv. Capitals are big cities. For example, uie aoo a ate e e s ide e haelo t l tie nne neefnat i e i ree n a , u t i r ier ita asentaees t e n sr rtht i a e fr r ass t rsnou hoot t s n t e c e r a t e e oe ia
- v. A cheap alternative toeit ea ooea t d e i s ne ev a r eltrat Ohnntenlfetine a i eh eao a l r ei e do ohls o er a seen ise as ae a enh u a a n e ine eteroe.al ae eoar o sn nen h ri e ø ne