Hiva Mohammadzadeh

# CS288 HW1 Report

I first increased the dimensions of the embedding from 512 and 1028 and it significantly improved my perplexity.

I used activation regularization and used a l2 regularization penalty and it lowered the perplexity in less epochs.

I also tried changing the learning rate just to see if that has positive effects: But the default (1e-3) was the best.

I also tried different batch sizes but didn't work.

Decided in order to increase efficiency, to keep track of the perplexity and only change it if it's less than the last one. So I had to change a lot of my code to make sure I recover models. Because I don't want a better perplexity to be overwritten by a worse perplexity. Added this to the regular model later as well because I was getting below the threshold when I was running 20 epochs.

Learning rate scheduling - added a learning rate scheduler that will decrease the learning rate by 0.2 at [10, 12, 14, 16, 18] epochs. Thought this might work because it decreases the learning rate as we get closer to the min and therefore it will help us not pass the min. This did help the accuracy a lot.

Ran for 1 epoch: We added a dropout of 0.5 after each LSTM layer. The dropout of 0.5 is bringing its perplexity from 253 to 263 which is already not helping. But since it was recommended to use 0.5 to get below 130, I thought I'll change it to see the result and it might help.

With the dropout of 0.4 on the model Lr = 1e-3 and weight decay of 1e-6 got 254 perplexity which is better than dropout of 0.5. So I thought I'd try decreasing it more and tried a dropout of 0.3 which gave me a perplexity of 243 which was better than 0.4 and better than without dropout . So, I decreased more to 0.1 which gave me 241 which was again better. So, therefore changing it to 0.1 helped my model for 2 epochs but ended up being worse when I ran the model for more than 2 epochs. So, I just stuck with 0.5 as my dropout.

Using all of this I was able to bring down my perplexity from 132 to 109 which is a large improvement.