Matthew Oakley

Deep Learning /Tensor Flow

Dr. Rivas

October 23rd, 2017

Hidden Units/seq length	25	50	75
32	1.772	1.825	1.826
64	1.535	1.623	1.565
128	1.345	1.419	1.364
256	1.171	1.212	1.178
512	1.056	0.991	0.835

Figure 1.

Increasing hidden units helps lower the perplexity which is also learning rate. This is because it is using more lsdm's to predict what the next word will be. If the hidden units is a low number and the sequence size is increased it makes the perplexity worse because there are not enough lsdm's for the size of the data. The sentences later on start with the word, "the" this is because many of the sentences in the data set start with the word, "the". This could train the system to always start with, "the" but when the sentence does not start with, "the" the accuracy

will be off. As seen in figure 1 increasing hidden units makes the perplexity lower because it uses more lsdm's.

Increasing the sequence length also helps lower the perplexity in the highest hidden units and this is because it uses more letters in each lsdm set. This gives it more training data to work with and makes it not encounter the problem with always starting with, "the". This is because it has more letters to work with and sees more complete sentences. The time also takes longer because each lsdm has to use more characters each epoch. Looking at figure 1 it can be seen that increasing sequence increases perplexity when there is not a high number of hidden units.