CSE 517 - Homework 5 - Reading Comprehension

Pooja Sethi - March 17, 2017

1.3 Train and Test CBOW Model

With the default hidden size of 50, I obtained the following results on the test set:
loss: 8.5229, acc1: 5.173%, acc2: 5.432%, EM: 3.285%, F1: 9.924%

We can see that the performance of the CBOW model is quite low, with an F1 score that is less than 10% and and EM score of less than 4%. CBOW does poorly because is not able to pick out long-range dependencies between context words. GRUs on the other hand are good at learning intra-sentence dependencies.

1.4 Gated Recurrent Units (GRUs)

Dropout 0.1:
loss: 5.7013, acc1: 22.203%, acc2: 26.064%, EM: 19.175%, F1: 29.956%

Dropout 0.3:
loss: 6.4201, acc1: 15.753%, acc2: 18.736%, EM: 12.023%, F1: 21.530%

Dropout 0.5:
loss: 6.8553, acc1: 13.778%, acc2: 16.323%, EM: 9.829%, F1: 19.153%

I found that in general increasing the dropout (and therefore lowering the input_keep_prob and output_keep_prob) causes the F1 and EM scores to dramatically decrease.

1.5 Attention

I added attention with a dropout rate of 0.1 on top of the GRUs in the previous question. I found that adding attention slightly increased the F1 score, but not as much as I was expecting it to.

loss: 5.8089, acc1: 22.203%, acc2: 25.669%, EM: 18.649%, F1: 30.086%