CS 6501 Natural Language Processing - Independent Project 3

Xin Nie, xn9vc

November 2018

1 Simple RNN LM Implmentation

2 Perplexity Implementation

xn9vc_perplexity.py

3 Perplexity Result

This section is run with epoch = 10, with embedding_dim = 32, hidden_dim = 32, num_layers = 1, lr = 0.2, training on one epoch takes about 3.5 minutes

- Training dataset: 540.863080348888
- Development dataset: 454.79046993621586
- Testing log probabilities: xn9vc-tst-logprob.txt

4 Stack LSTM Implementation

This section is run with epoch = 10, with embedding_dim = 32, hidden_dim = 32, lr = 0.2
Tried n = {1, 2, 3}, and the better value of n is: 3
Training on one epoch when n = 2 is about 5 minutes, when n = 3 is about 7 minutes

- Training dataset perplexity: 465.6150334646841
- Development dataset perplexity: 401.39883728281416

5 Optimization

This section is run with epoch = 10, with embedding_dim = 32, hidden_dim = 32, num_layers = 1, lr = 0.2
Tried SGD with momentum and AdaGrad method, the better model is: AdaGrad

- Training dataset perplexity: 443.1552813518375
- Development dataset perplexity: 457.61034520239747

6 Model Size

This section is run with epoch = 10, with embedding_dim = 32, hidden_dim = 32, num_layers = 1, lr = 0.2
Tried the input/hidden dimension = {32, 64, 128, 256}, tried different combinations, the better is: {256, 256}

- Training dataset perplexity: 312.3689342676171
- Development dataset perplexity: 329.26801956279