CS310 NLP Assignment 2 -- Word2Vec

钟志源 12110517

Step 3

Example loss change:

```
Iteration: 0
| epoch   0 |   500/ 2485 batches | loss    2.461
| epoch   0 |  1000/ 2485 batches | loss    3.130
| epoch   0 |  1500/ 2485 batches | loss    1.971
| epoch   0 |  2000/ 2485 batches | loss    1.936
current lr: [0.02200000000000002]
iteration 0 Loss: 2.9090666740712985



Iteration: 1
| epoch   1 |   500/ 2485 batches | loss    2.077
| epoch   1 |  1000/ 2485 batches | loss    2.917
| epoch   1 |  1500/ 2485 batches | loss    1.538
| epoch   1 |  2000/ 2485 batches | loss    2.763
current lr: [0.01936000000000002]
iteration 1 Loss: 2.33308997096669997



Iteration: 2
| epoch   2 |   500/ 2485 batches | loss    2.101
| epoch   2 |  1000/ 2485 batches | loss    2.220
| epoch   2 |  1500/ 2485 batches | loss    2.389
| epoch   2 |  2000/ 2485 batches | loss    1.564
current lr: [0.0170368]
iteration 2 Loss: 1.9663208678934416



Iteration: 3
| epoch   3 |   500/ 2485 batches | loss    1.470
```

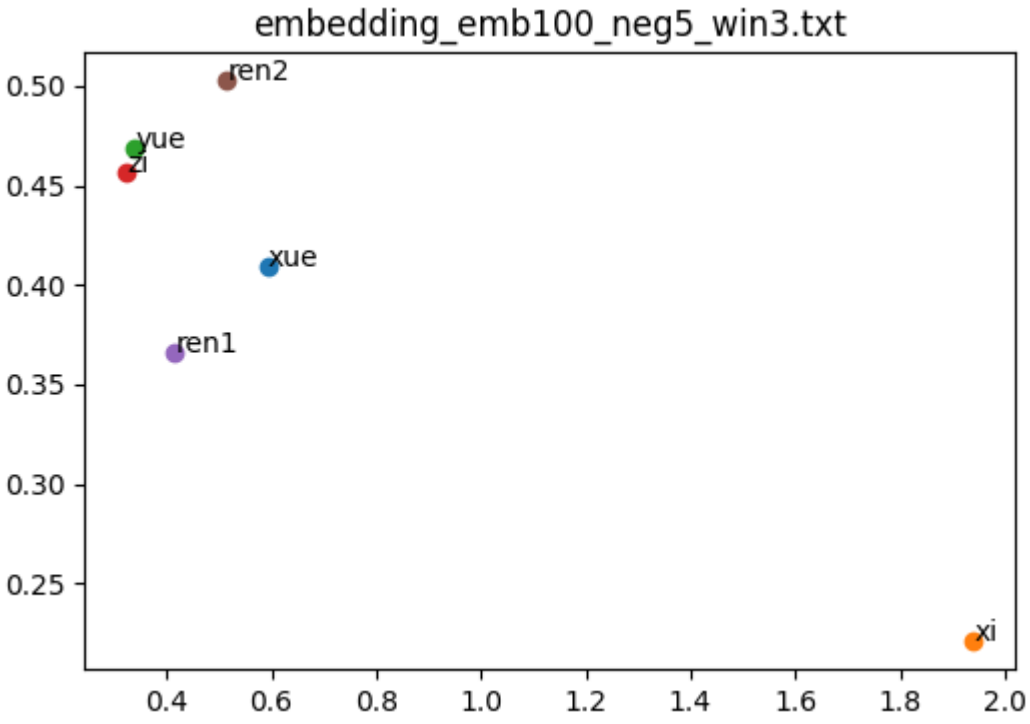
```
| epoch   3 |  1000/ 2485 batches | loss    1.761
| epoch   3 |  1500/ 2485 batches | loss    2.246
| epoch   3 |  2000/ 2485 batches | loss    2.454
current lr: [0.014992384000000001]
iteration 3 Loss: 1.7799131879144512



Iteration: 4
| epoch   4 |   500/ 2485 batches | loss    1.612
| epoch   4 |  1000/ 2485 batches | loss    1.531
| epoch   4 |  1500/ 2485 batches | loss    2.054
| epoch   4 |  2000/ 2485 batches | loss    1.846
current lr: [0.013193297920000001]
iteration 4 Loss: 1.6747972374710758
```

Usually when the epoch is around 10-15, the loss stops decreasing significantly. So I set the epoch to 15.

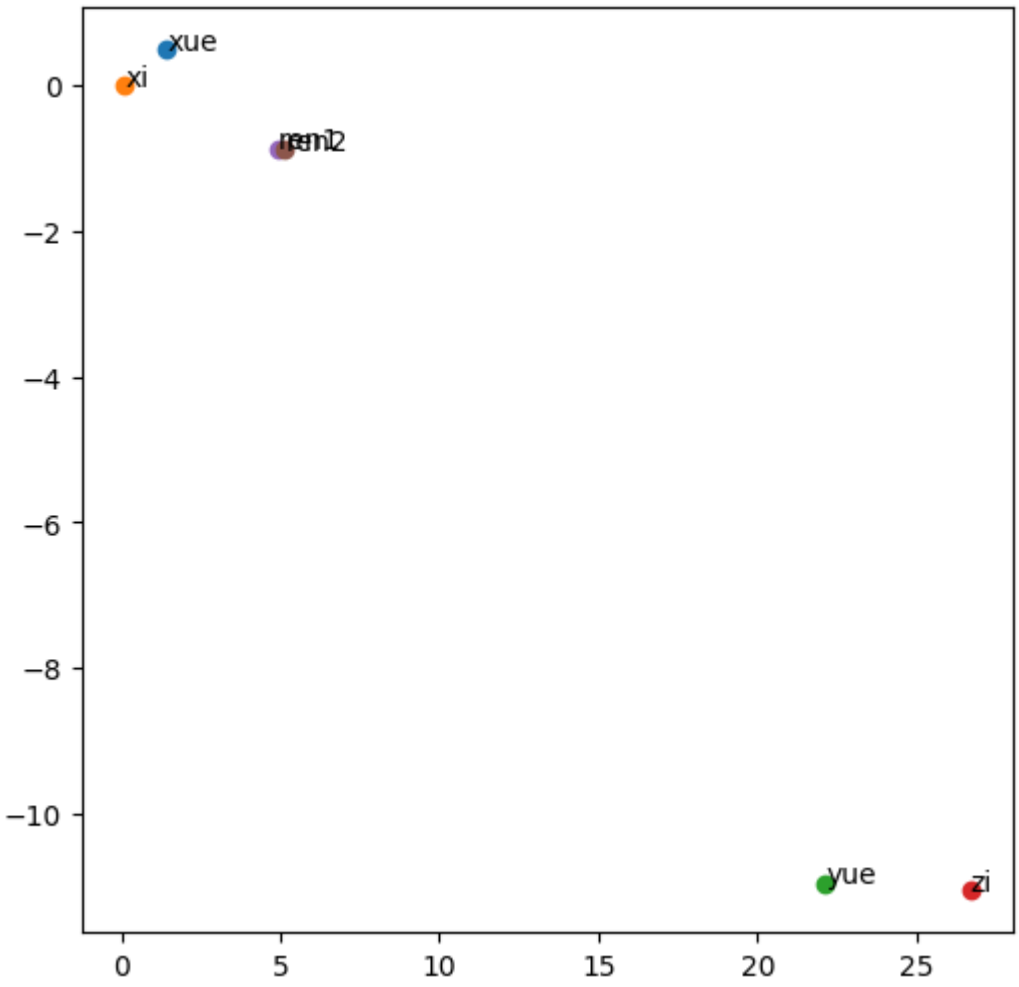
Step 5

Comparison with LSA embedding

Embedding from skipgram model:



Embedding from LSA model:



The 子 and 曰 are close to each other in skipgram and lsa. But the 学 and 习 are distant in skipgram, while close in lsa. Word meaning is pretty tricky:)