Applied Deep Learning Assignment 1

Tung-Chun, Chiang R05922027

November 9, 2016

1 Embedding Result

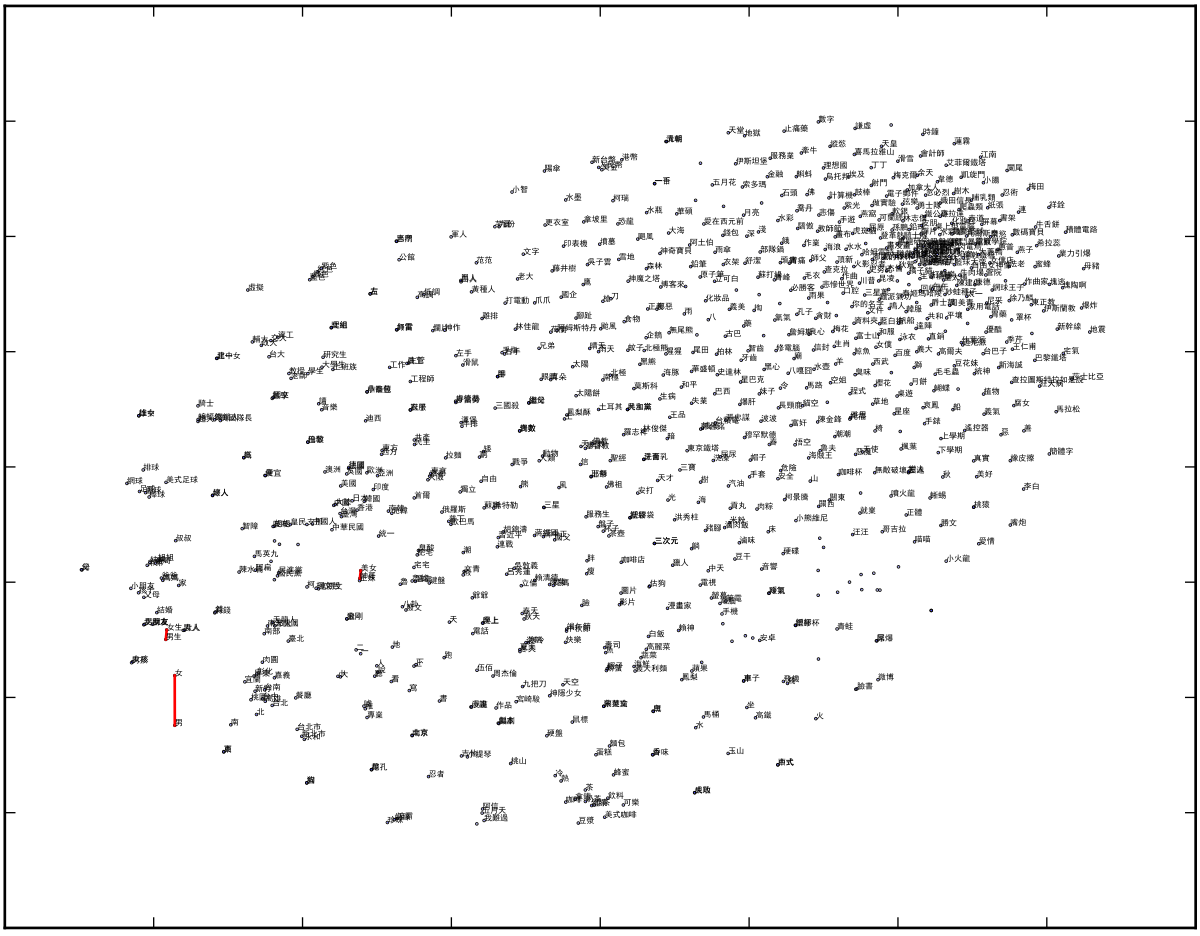

Figure 1: GloVe word embeddings for ptt corpus.

Figure 1 is generate by tSNE method. tSNE uses PCA Dimensionality Reduction and reduces dimension to 2. This allows embeddings to be shown on an image. The embeddings in Figure 1 are trained with ptt corpus and GloVe algorithm $(n\_epochs = 80, n\_dims = 150)$.

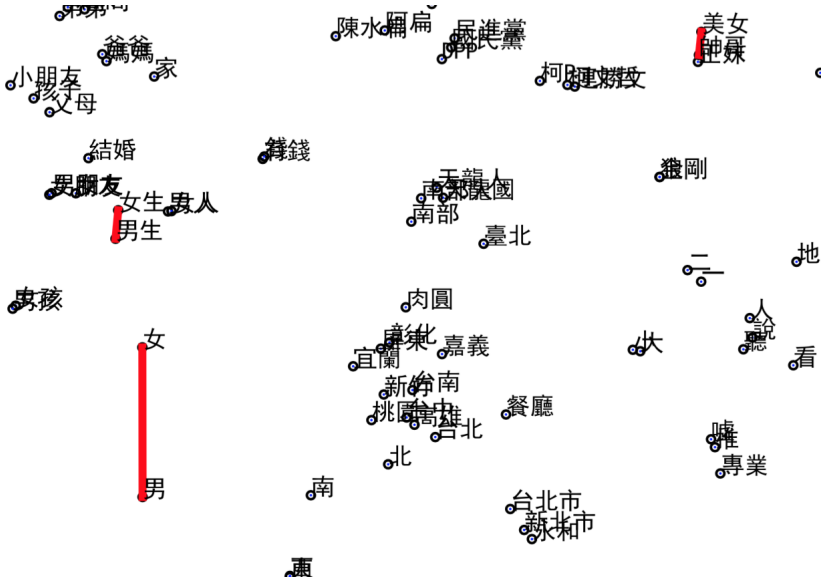

Figure 2: Word pairs with male and female relation.

Figure 2 shows some word pairs with the same relation: male and female. You can see that the directions and length are almost the same, and it means this model actually learns some relationships between words.

2 Training Records

- Remember to dump model! So that you can train a new one with pre-trained model as initial.
- Learning rate should decrease during training.
- Properly name you model and vector, or maybe overwriting happens.
- Training with threading(word2vec_optimized.py) can save a lot of time.
- Training with GPU can save a lot of time.
- Implement your own model as possible.
- Tracing others' codes can help a lot.