**Word2vec Analogical Result Report**

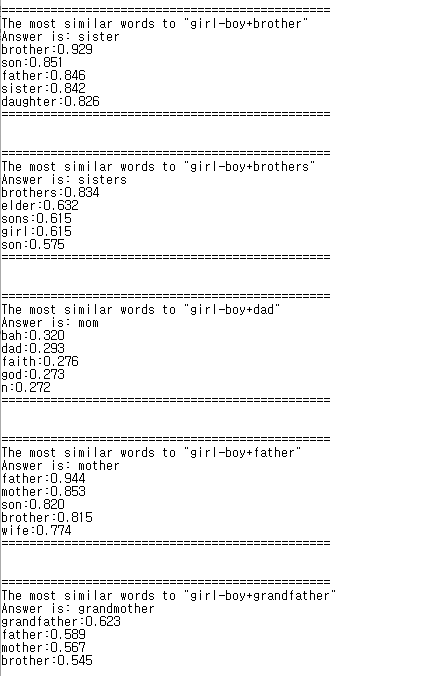
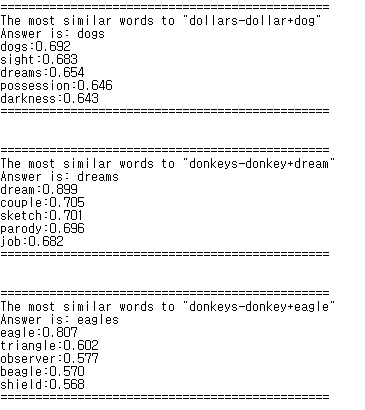
2016320134 컴퓨터학과 강동식

Word2vec에 Hierarchical Softmax와 Negative Sampling를 이용하여 CBOW를 통해 학습한 결과들이다. Skip-gram은 학습 시간이 길어 결과에 추가하지 못했다. Dimension은 300, window size는 5를 통해 학습하였고, epoch는 1번 학습하였다.

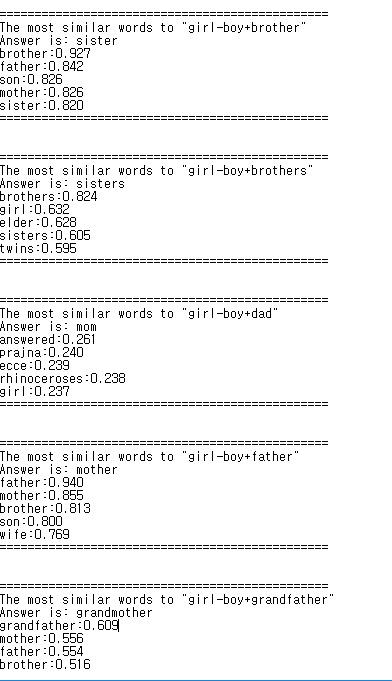
CBOW보다 Skip-gram을 이용하여 학습한 것이 window size인 10만큼 더 학습한 것이기 때문에 더 좋은 결과가 나타날 것으로 예상된다. Subsampling을 이용하여 학습한 모델은 하지 않은 모델보다 약간 더 좋은 학습 결과를 나타난 것으로 보이고 Negative sampling이 hierarchical softmax보다 좋은 학습 결과와 빠른 실행 속도를 보이는 것으로 관찰되었다.

학습은 word2vec\_subsample.py를 이용해 학습하였고, 이후에 epoch마다 subsampling을 하도록 수정하는 파일인 word2vec.py로 코드를 수정하였다.

CBOW\_5\_full\_subsampled

CBOW\_5\_full\_notsubsampled



CBOW\_0\_full\_subsampled

