

이태일 pca는 비선형성이 없고, word2vec

은 비선형성이 있습니다. 100차원을 2차원으로 pca로 줄일 시 선형성에만 의존하여 정보 손실이 더 있을 수도 있지만 저라면 100차원으로 word2vec 을 만들고, 시각화할 때만 pca로 2차원으로 만들어 사용할 것 같습니다. 왜냐하면 100차원 벡터가 2차원 보다 훨씬 많은 정보를 가지고 있을 것으로 사료되기 때문에 실제 사용 시에는 100차원, 시각화 시에만 따로 2차원으로 사용하는 게 맞는 것 같습니다. 참고로, Pca대신 오토인코더를 사용하시면 비선형으로 100차원을 2차원으로 축소도 가능하십니다.