Roll: 2019121004
Name: Avani Gupta

Advance NLP Assignment 2

ELMo

Implemented in python
Libraries used: Keras, TensorFlow
Encoding: Word encoding

Preprocessing:

- Sentences were taken from given dataset trans files.
- <bos> <eos> added to begin of sentence and end of sentence.
- Lowercase words
- remove special chars
- Train-Val split: 90:10

Vocab:

length: 26956

Training:

Trained for 5 epochs: with Val loss becoming 3.913 in 5th epoch and staying constant then.

Distances between word pairs

Notations
ed: Euclidean Distance, cs: Cosine Similarity

pretty-target: ed: 33.5099983215332 cs: 0.29

pretty-nice: ed: 27.690000534057617 cs: 0.36

pretty-days: ed: 32.119998931884766 cs: 0.26

pretty-growing: ed: 28.25 cs: 0.31

pretty-dress: ed: 26.420000076293945 cs: 0.31

pretty-target: ed: 33.5099983215332 cs: 0.29

last-dress: ed: 28.920000076293945 cs: 0.33

last-growing: ed: 35.849998474121094 cs: 0.46

target-fast: ed: 35.22999954223633 cs: 0.33

target-nice: ed: 43.59000015258789 cs: 0.54

Discussion:
The distance of word pairs pretty-nice is smaller than pretty-days which means nice is closer to pretty than days in the ELMo embeddings.

References:
Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*.