Files List – Appendix

1. QA-LSTM: this folder contains all the code and the vectorized dataset to train and test this model. The notebook is titled QA-LSTM.ipynb.
2. BERT Fine-Tuning: contains the tokens for training and validation sets to fine tune the BERT base, uncased model. There is also a model folder with saved models in case one would like to load that to perform predictions. Lastly, the code for training and testing is also included in the notebook BERT\_QA\_Pairs.ipynb.
3. BERT Embeddings: contains the saved tensor and numpy embeddings for q and a used to perform cosine similarity and training other models. Code is included in BERT\_Sentence\_Embeddings.ipynb.
4. Two additional data csv files are also included, with slight modifications to the original Train.csv file; Train\_w\_cosine\_sim.csv and Train\_RemovedNAN\_AddedNormScore.csv.