**Homework 12:**

**About the Corpus**

The Stanford Natural Language Inference (SNLI) corpus (version 1.0) is a collection of 570k human-written English sentence pairs manually labeled for balanced classification with the labels “entailment”, “contradiction”, and “neutral”. We aim for it to serve both as a benchmark for evaluating representational systems for text, especially including those induced by representation-learning methods, as well as a resource for developing NLP models of any kind.

The following paper introduces the corpus in detail. If you use the corpus in published work, please cite it:

Samuel R. Bowman, Gabor Angeli, Christopher Potts, and Christopher D. Manning. 2015. A large annotated corpus for learning natural language inference. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP). [pdf] [bib]

In the demo I gave during the class time, I built a Transformer encoder using “building blocks” from Keras in order to classify the movie reviews into positive or negative sentiment. In this homework, different from the demo where I built a Transformer encoder using building blocks, I want you to use an existing a pretrained BERT model (the “bert-base-uncased” version), and then add a BiLSTM layer (with average pooling) to classify whether a pair of sentences entail, contradict, or have a neutral relation.

Please use only the first 10,000 samples from the training data to train the BiLSTM layer. The pretrained BERT freezes. You don’t need to train. Please set the batch size 32, epochs 1, and the maximum input sequence length 128.

As the result, please just show the one epochs’ training loss and accuracy as well as validation loss and accuracy.