Assignment 7: Embeddings, Recurrent Neural Networks, and Sequences (Part 2)

Start Assignment

Due Tuesday by 5:59pm **Points** 6 **Submitting** a text entry box or a file upload

Question 7 (1 point)

Assuming the activations of the layer above consists of a sequence of embeddings with (d_model=) 1024 floating-point values per embedding, and the desired dimensionality of the feed-forward expansion is (d_ff=) 4096, what would be the dimensions of the weight matrices for a transformer block?

Model 7 (5 points)

Please navigate to the following URL to accept the invitation for this Kaggle task:

https://www.kaggle.com/t/26ce7f41271e478d9ee0814485c7ee09

Activate the conda environment on your VM:

conda activate py37_tensorflow

Install the library for the Recognition and Organization of Speech and Audio (librosa):

pip install librosa

Download the data and create the tensors for the ".wav" files:

kaggle competitions download ml530-2021-sp-speech

wget https://www.cross-entropy.net/ML530/speech-tensors.py.txt

python speech-tensors.py.txt

Run the sample training script (requires both transformer.py and speech-train.py.txt):

wget https://www.cross-entropy.net/ML530/transformer.py

wget https://www.cross-entropy.net/ML530/speech-train.py.txt

python speech-train.py.txt

For a super stretch goal, consider augmenting your training data with a bit of noise. To avoid filling the disk, you'll probably want to implement a https://keras.io/api/utils/python_utils/#sequence-class :

https://www.cross-entropy.net/ML530/add_noise.py.txt

https://www.cross-entropy.net/ML530/noise.zip