HW3 Bootcamp: P2
Data and Task

- **Features:**
  - Same as hw1p2 (array of utterances with 40D timesteps)
  - No P2 multiple choice

- **Labels:**
  - Order? Alignment?
  - List of lists - cast to numpy array

- **Must generate sequences of phonemes**
  - 41 phonemes and 1 blank character

- **Loss:** CTCLoss

- **Metric:** mean Levenshtein distance
  - Can import
Modeling and RNN+Variants

- Can use other types of layers
  - Hint: Convolutional Layers
  - Input shape: (batch, in_channel, length)

- RNN, LSTM, GRU, etc.
  - Capture sequential dependencies
  - Input shape: (length, batch, feature) or (batch, length, feature)

- No attention

http://colah.github.io/posts/2015-08-Understanding-LSTMs/
Batch of Variable Length Inputs: Padding

- HW1, HW2: equal length inputs
- HW3: variable length sequences
- Method 1: Pad
  - Inefficient with space
- Method 2: Packing
Batch of Variable Length Inputs: Packed Sequence

Figure 2: List of tensors we want to pack

Figure 3: First we sort the list in a descending order based on number of timesteps in each

Figure 4: Final Packed 2d Tensor

Reference: Fall 20 HW3P1 Writeup
Packed Sequences

- **pad_sequence()**
  - Pads to equal length for batching

- **pack_padded_sequence()**
  - Packs batch of padded sequences
  - Requires sequences + sequence lengths

- **X = pad_packed_sequence()**
  - Unpacks back to a batch of padded sequences
  - Outputs sequences + sequence lengths

- **Collate Function**
  - Dataloader argument
  - Helpful when altering data for batch

Figure 4: Final Packed 2d Tensor
Output Processing

Output = probability distribution at each timestep

- Order Synchronous, not time synchronous
- Greedy Search
- Beam Search
  - Marginal performance improvement
  - Linked in the writeup
- Feel free to use your own!
Advice

- Watch lecture and recitations
- Read the entire write-up and understand the problem before starting
- Look at the example submission
- Check tensor shapes
  - batch_first=True
- Use what you’ve learned from previous P2s
Some Helpful References

- Homework_3_1.pdf (cmu.edu) - http://deeplearning.cs.cmu.edu/F20/document/homework/Homework_3_1.pdf
- http://colah.github.io/posts/2015-08-Understanding-LSTMs/