In-Class Assignment (e.g. Homework # 4)

Due Date: Thursday, November 21, 2019 by 11:59PM
Total Points: 100

This assignment introduces you to one speech recognition toolkit and will let you experiment with training and testing a speech recognition system. You will be using Kaldi, which is a popular toolkit, to train a speech recognition system for the TIMIT speech corpus. **For the purposes of this assignment, please work individually**. However, feel free to ask basic questions to the instructor and classmates, but complete the assignment on your own.

You must submit a typed report that introduces the problem and discusses all results. Your report should be submitted as a single pdf document through Canvas. Also submit the modified run.sh script.

All assignments must be submitted on time to receive credit. No late work will be accepted, unless you have a prior arrangement with the instructor.

Question 1. [100 POINTS]

1. **Kaldi Installation**:
   - Download Kaldi from: https://github.com/kaldi-asr/kaldi.git
   - Go to /tools and follow INSTALL instructions there
   - Go to /src and follow INSTALL instructions there
2. **Download TIMIT dataset**
   - The dataset is located at: Canvas → Box Course Folders. Go to COURSE FILES → Speech Data → TIMIT_full.zip
   - Download and unzip the file, then store the contents locally. Be sure to make note of the path to the dataset
3. **Kaldi with TIMIT dataset**: Kaldi provides example scripts for many publicly available datasets. The scripts can be found in egs/ directory. This documentation is about how to use Kaldi with TIMIT dataset with example scripts in egs/timit/s5.
   - Before running the scripts, a couple things need to be completed:
     - Go to kaldi/tools/ then run: extras/install_irstlm.sh
     - Add source ../../../tools/extras/env.sh to path.sh in kaldi/egs/timit/s5
   - By default, the scripts assume the system has a queue system. To run them on the local machine, change all instances "queue.pl" to run.pl in cmd.sh
   - To enable GPU code, open steps/nnet2/train_tanh.sh and change num_threads=16 to num_threads=1
   - The script run.sh should be executed for training and testing.
     - The script contains training and testing code for different speech recognition models, including DNN, sGMM and some language models.

- Code under each comment is for the specific system, such as
  echo=====================================
  echo " DNN Hybrid Training & Decoding "
  echo=====================================,

  * Comment out these sections: "System Combination (DNN+SGMM)" and "DNN Hybrid Training & Decoding (Karel's recipe)"
  * Set timit=[path_to_dataset_is] in run.sh
  * Parameters can be added to this line steps/nnet2/train_tanh.sh in the following format: Training: steps/nnet2/train_tanh.sh –name_of_parameter parameter_value

4. **Run Kaldi**: Enter this in the command line: ./run.sh., to run the Kaldi script that trains and tests different ASR systems. Discuss the results in your report. Be sure to compare how the different approaches perform.

5. **Modify and Re-run Kaldi**: Modify the script in the following way:

   - Modify 'DNN Hybrid Training & Decoding' so that the DNN uses 1, 3, and 4 hidden layers. Meaning you need to run the script three separate times
   - In your report, specifically discuss how the number of hidden layers impacts performance.