Lab 5: Group work on projects

The goal of this lab is for you to make progress on your project, together as a group. You'll set goals and work towards them, and report what you got done, challenges you faced, and subsequent plans.

Group name: How to turn this on.

Group members present in lab today: Yi-Ting Yeh, Ting-Rui Chiang

1: Plan

1. What is your plan for today, and this week?
   Today: Set up the development environment and finish the data preprocessing scripts.
   This week: 1) Understand how to customize models in ESPNet. 2) Finish the implementation of attention head pruning and layer pruning. 3) Have initial results if possible.
2. How will each group member contribute towards this plan?
   Yi-Ting: Layer pruning.
   Ting-Rui: Head pruning.

2: Execution

1. What have you achieved today / this week? Was this more than you had planned to get done? If so, what do you think worked well?

   We set up the development environment and write the data preprocessing scripts for our experiments. We dig into the source code of ESPNet2 and understand its different training stages. We modified the samples training ASR model scripts on the LibriSpeech dataset to meet our experiment settings and use pretrained models. We finish the implementation of attention head pruning and layer pruning which includes LayerDrop and intermediate CTC.

2. Was there anything you had hoped to achieve, but did not? What happened? How did you work to resolve these challenges?

   For layer pruning, we are still training the first model which fine-tunes the pretrained Transformer ASR model with LayerDrop and intermediate CTC. It is because the implementations of LayerDrop and intermediate CTC are a little tricky. The two features were originally implemented in ESPNet 1 but not ported to ESPNet 2, which is the latest

version we use. Therefore, we need to directly modify the code in the installed library to re-implement the LayerDrop and intermediate CTC. It takes more time than we expected so we haven't finished training the first model.

3. What were the contributions of each group member towards all of the above?
   Yi-Ting: Layer pruning.
   Ting-Rui: Head pruning.

3: Next steps

1. Are you making sufficient progress towards completing your final project? Explain why or why not. If not, please report how you plan to change the scope and/or focus of your project accordingly.

   Yes, we are on the right track and have sufficient progress towards completing the project. We have finished the hardest part which is understanding the code structure of the ESPNet and know how to modify it to train the models. Now we can spend more time on experimenting with different types of models.

2. Based on your work today / this week, and your answer to (1), what are your group's planned next steps?

   We want to firstly finish the training of layer pruning models and have its initial results. If the results don't make sense, we will need to identify the issues and fix bugs.

   After we have initial results of layer pruning, we will apply head pruning over it. We will also inspect the performance when a neural-based language model is not used. If a neural-based language model is necessary for low WER, we may need to prune the language model too.

3. How will each group member contribute towards those steps?

   Yi-Ting: Finish the training of layer pruning models and fix any potential problems. Test it on our device.
   Ting-Rui: Apply head pruning on the layer pruned model, and also inspect the performance when neural-based language model is not used.