**Jinnan Li**

**Topic proposal**

**Hybrid CTC/Attention Architecture for End-to-End Speech Recognition**

**Topic category**: end-to-end ASR

**Goals and motivation**: We will reproduce works in the paper (Hybrid CTC/Attention Architecture for End-to-End Speech Recognition). To address problems of CTC (e.g. independence condition) and attention architecture (e.g. difficulty to track the alignment). The paper proposed a new architecture combining both CTC and attention architecture. During training, we propose a multiobjective learning method by attaching a CTC objective to an attention-based encoder network as a regularization. During decoding, we propose a joint decoding approach, which combines both attention-based and CTC scores in a rescoring/one-pass beam search algorithm to eliminate the irregular alignments. We plan to implement this new architecture and compare it with the baseline models (CTC, attention architecture, and maybe HHM/DNN) using the provided database.

**Resources**: S. Watanabe, T. Hori, S. Kim , J. R. Hershey, and T. Hayashi, "Hybrid CTC/Attention Architecture for End-to-End Speech Recognition," (Links to an external site.) IEEE JSTSP 2017.

**Data and experiments**: using the provided database on canvas and maybe databases used in the paper (e.g. WSJ and CHiME-4) to train the new proposed and baseline models.

**Additional computing resources**: Maybe an HMM/DNN model in GitHub (stated later).

**Timeline**: Testing baseline models on week 8 and testing the hybrid model on week 9.