The input is the a audio data stream and the ground truth, formatted as a numpy matrix (each row is the label for a second audio)

1. To extract features: call A_extractFeature_combined.py, then A_contextWin_combined.py

2. To train dnn: A_DL_combined_target.py

3. To train cnn: A_cnn.py

4. For Demo(process data second by second): A_Demo_Dog_alarm_doorbell.py