

## Topic : English & Korean Automatic Speech Recognition for Understanding Voice-Commands

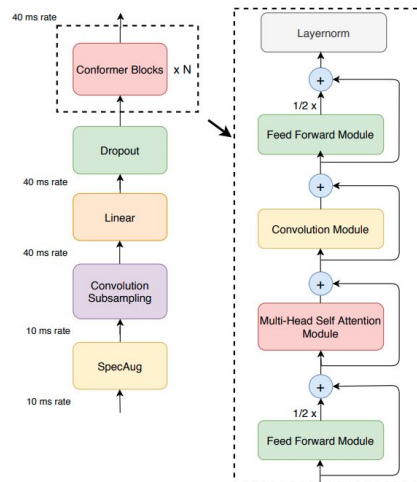
### [Current Task]

- Train ASR model w/ Korean & English speech dataset



### [This Week]

- Study ROS : Writing nodes, packages
- Related Works : recent ASR models
  - Conformer (Gulati et al. 2020), ContextNet, SqueezeFormer, etc.
- Dataset Gathering
  - LibriSpeech : 1000+ hrs, audio-books, English
  - KsponSpeech : ~1000 hrs, conversations, Korean ([AIHub.or.kr](http://AIHub.or.kr))
- KoSpeech / OpenSpeech Library
  - Supports KsponSpeech (+ preprocessing)
  - 20+ latest SOTA models (Conformer, etc.)
  - (Problem) KsponSpeech dataset files not readable (corrupt?)
  - (Problem) Other bugs & issues -> fixing GitHub code as needed



### [Next Week]

- Fix dataset issues
- Try KoSpeech model training w/ Korean dataset (*KsponSpeech*)
- Try other libraries : Nvidia NeMO, torchaudio, etc.

```
# audio_file_name [tab] 한글 전사 [tab] 숫자 전사
$DATA_PATH/audio_01.pcm 점심 맛있게 먹었니 320 1 3 24 56 88 3 98 124 76
...
$DATA_PATH/audio_02.pcm 너 뭐 강남 자주 가? 51 3 42 3 243 197 3 47 77 3 9 15
```