

가가가 가? 시리야 이거 표준어로 번역해줘

붉은악마의자식 🐼 권보영 서지영 한채현

Motivation

최근 유튜브 <경상도 호소인>, 드라마 <내 남편과 결혼해줘> 등 지역별 사투리에 대한 관심이 높아지고 있다. 따라서 사투리 발화를 인식하여 표준어 발화 대응쌍을 생성해주는 모델을 만들어보고자 하였다.

Method

모델은 크게 세 파트로 나눌 수 있다.

- Ko-Speech (Speech-to-Text)

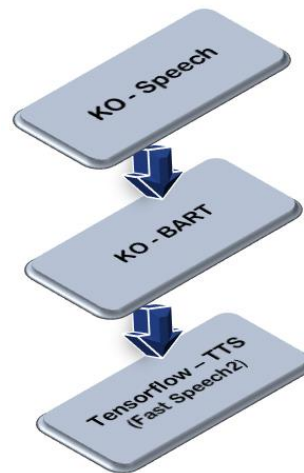
DeepSpeech2, Speech Transformer 등 다양한 알고리즘 기반의 STT 모델로 한국어만 지원

- KoBART (Text-to-Text)

BART = BERT + GPT, KoBART는 한국어 데이터로 학습된 encoder-decoder 언어 모델

- Tensorflow-TTS (Text-to-Speech)

실시간 음성 합성 아키텍처로, 본 프로젝트에서는 FastSpeech 2를 기반으로 한 모델 사용



Experiments

Ko-Speech 학습 (w/Kspon)

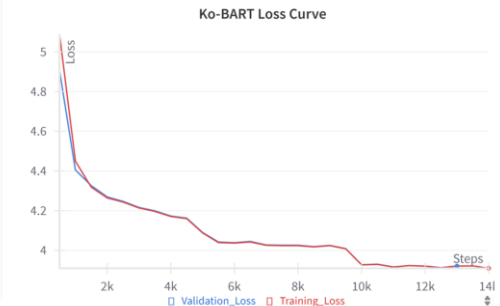
평균 CER : 0.38

79) [DKSR20000890_1_1_86.wav]
[transcript] 어 어떤 사람 만났을 때 어 그런 생각이 들면은
[prediction] 어 어떤 사람 만났을 때 어 그런 생각이 들면은
Ncor: 9 Nsub: 0 Ndel: 0 Nins: 0
Lev distance: 0 length: 18
WER: 0.0 CER: 0.0

323) [DKSR20000891_1_1_88.wav]
[transcript] 일곱 시까지 항상 만화를 많이 했잖아
[prediction] 일곱시까지 항상 만화를 많이 했잖아.
Ncor: 3 Nsub: 2 Ndel: 1 Nins: 0
Lev distance: 1 length: 15
WER: 0.5 CER: 0.06666666666666667

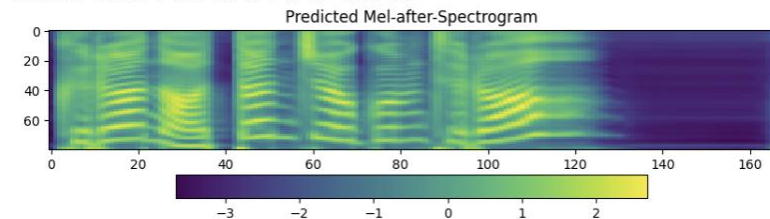
KoBART 학습

(w/ 한국어 방언 발화(경상도))



Tensorflow-TTS를 포함한 전체 파이프라인 시연 결과

Ko-Speech 결과(음성인식) : 그런 어떤 좀 들어?
Translated Standard (표준어로 출력) : 그런 어떤 조금 들어?



0:00 / 0:01