

5가지 API 비교

작성일 2022.08.03

작성자 윤수연

사용 음성파일: 2022.07.11 인하관청의 회의 음성

원본 음성 길이: 30분 50초

전처리: 음성 자르기

음성 전처리 규칙

- 한 사람의 발화만을 담을 것
- 음성의 길이가 10초를 넘어가지 않을 것
- sampling rate: 16000 Hz
- file type: wav file

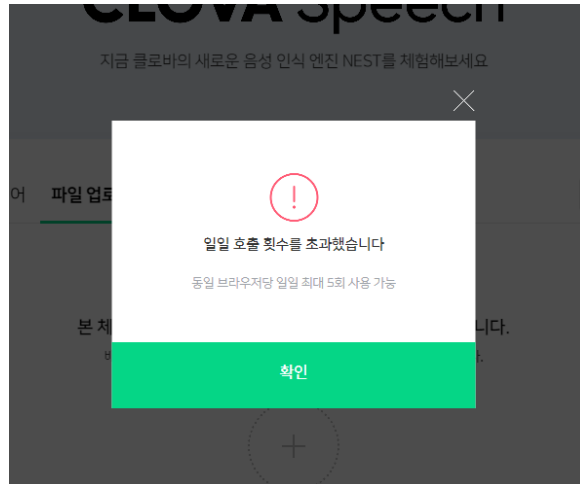
전처리 결과물(10)

1. 용량이 어느 정도 되나요 혹시 학교 서버가?(2)
2. 학교서버요?(0)
3. 네 저장장치 용량(1)
4. 적어도 테라바이트 이상은 돼야죠. 왜냐하면은...(2)
5. 저희 연구실 서버 말씀하시는 건가요 혹시?(2)
6. 연구실 서버요(0)
7. 수어를 선호하는데 매체가 없어서 말로 해야되거나 글로 해야되거나 할 수 있으니까...(5)
8. 이거 두 개 다 일단 알고 싶어요ㅎ(2)
9. (쓰읍)시범으로 저희 그 저기 GUI 외주 줄 때 작성한 문건 있잖아요 그거 그냥 그대로 보여드리고(6)
10. 우선 저희 그거 공유해볼까요 오늘 멘토링(3)

비교 실험

1. 네이버 : [CLOVA Speech](#)

- 임의로 문장을 구분하여 끊어줌
- sampling rate 16,000Hz로 하면 비교적 빠름
- 일일 호출 횟수: 5회 → 3일에 걸친 실험



1. 용량이 어느 정도 되나요. 혹시 학교 서버가
- 2.
3. 네 저장장치 호장 용량
4. 적어도 테라바이트 이상은 돼야죠 왜냐하면
5. 저희 연구실 서버 말씀하시는 건가요 혹시
6. 연구실 서버예요.
7. 수호를 선호하는데 매체가 없어서 말로 해야 되거나 글로 해야 되거나 할 수 있으니까요.
8. 이거 두 개 다 일단 알고 싶어요.
9. 지원으로 그냥 저희 저 gi 외주주 때 작성한 문건 있잖아요. 그거 그냥 그대로 보여드리고
10. 우선 저희 그거 공유해볼까요. 오늘 멘토링

2. ETRI API: [AI API.DATA](#)

- REST API
- 입력조건: postman으로 sr 16000 16 모노 wav, POST타입으로 첨부



POSTMAN

-
- [변환사이트](#), Audacity로 input data 조건 충족
- 딜레이 거의 없음, 측정하지 않았지만 1초 미만

1. 용량이 어느 정도 되나요, 혹시 학교 서버가?
2. 수 없나요
3. 네, 저성장 지우장, 용량.
4. 적어도 테너바이트 이상은 돼야죠, 왜냐하면
5. 연구실서 말씀하시는 건가요, 혹시?
6. 연구결사부예요.
7. 수어를 선호하는데 매체가 없어서 말로 해야 되거나 근로해야 되거나 할 수 있으니까요.
8. 이거 두 개 다 일단 알고 싶어요.
9. 지원으로 그냥 저희 기와 외주 줄 때 작성한 문건 있잖아요. 그거 그대로 보여드리고.
10. 우선 저희 그거 공유해볼까요 오늘 멘토리

최종 결과

- [ysy2000의 github](#) 의 코드를 이용하여 CER계산

	CER 평균	CER 0점 개수
CLOVA	0.211982	4
ETRI	0.287763	2