

what is training set?

validation set

test set?

레이블이 과도하게 많을 경우 정확도가 떨어지는 일이 발생할 수 있으므로, 초기 모델에 입력 후 정확도를 비교해 {편견, 혐오, 그 외}로 할지 {편견, 성편견, 혐오, 불편한 발언, 그 외}로 할지 결정할 예정이다. - 얼마나 차이가 있는가 궁금

데이터셋 분할 ? 데이터 70, 검증 20, 테스트 10

1. 데이터 파이프라인
2. Model API 서버 작업
3. 프레임워크는 Flask, FastAPI
4. GitLab 사용이 더 좋을 것 같다.
5. <https://velog.io/@dogyun-k/RestTemplate로-다른-서버와-통신하기#개요>
6. <https://artist-developer.tistory.com/24> 데브옵스

그냥 간단하게 혐오에 대해서 뭘 더 추가할 필요가 없이 이 사람의 답글이 편견이 존재하는지 혐오 표현이 존재하는지 판단해주는 프로그램이라고 생각했어요.