

10월 및 대회 회고록

전체 돌아보기 :

대회 마감으로 최종 점수 상향에 집중.

대회 종료 후 소감 및 분석은 하단의 대표 1인 PDF 참고 부탁드립니다.

https://prod-files-secure.s3.us-west-2.amazonaws.com/8fb7aff8-a137-4ee9-b106-21abbfc59051/a027c454-6329-4ef2-a0c3-b2c5cd536aa7/CommonLit_Portfolio.pdf

9월 27일

- **이성준님** →
BERT 모델에 대해 한국어로 설명된 노트북을 공부해 보았고, BERT에 대해 완전히 이해를 함.
- **한찬혁님** →
다른 사람들이 공유한 Public 노트북을 참고하며 Submit 해보았음.
- **임승준님** →
debertav3base + LGBM 코드 등, 대회에 공유된 좋은 코드들 소개
bert-base 모델을 활용한 코드 필사 중
- **이희상님** →
MobileBERT를 제대로 구현하기 위해 Teacher모델 구현 완료하였으나, 점수가 제대로 나오지 않아 당황하였음.
Teacher Model에 deberta를 활용하였으나, 다른 모델로도 해볼 예정.

[EN/한/中] MobileBERT implementation

Explore and run machine learning code with Kaggle Notebooks | Using data from multiple data sources

[k https://www.kaggle.com/code/jasonheesanglee/en-mobilebert-implementation](https://www.kaggle.com/code/jasonheesanglee/en-mobilebert-implementation)



- **배나영님** →

병가로 인한 불참

9월 27일 ~ 10월 3일

- 해피 추석!!

이지만 개인적으로 시도해보고자 하는게 있으면 해보기!

10월 4일

- **배나영님** →

현재까지 작업하던 노트북 고도화 중

- **이희상님** →

추석 기간에 섭밋을 하긴 했으나... 개인 일정관계로 숫자 놀음만 하다가 끝냄.
대회 마지막이라 T5에 집중하고 있음

- **이성준님** →

추석기간에 가족 일정으로 특별히 무언가 해볼 기회가 없었음.

- **한찬혁님** →

Submit했던 노트북 중 가장 좋은 점수의 노트북에서 파라미터 조정하여 CV 점수 확인
하여 볼 예정

- **임승준님** →

병가로 인한 불참

10월 4일 ~ 10월 10일

- **이희상님** →
대회 마무리까지 T5 및 타 기법 활용하여 최대한 점수 올려보기
- **이성준님** →
마지막으로 BERT 모델 노트북 고도화 예정
- **한찬혁님** →
Submit 했던 노트북 중 가장 좋은 점수의 노트북에서 파라미터 조정하여 CV 점수 확인하여 볼 예정
- **임승준님** →
마지막으로 여태 배웠던 내용 참고하여 섭밋해 볼 예정

10월 11일

명일 오전까지 섭밋해야하여 어떤 노트북을 최종 선택할 지 마지막까지 고민하였음.

- ▼ 당시 상황의 분위기 전달을 위하여 아래 슬랙 채널 채팅 공유합니다.



이희상_오프라인 🇰🇷 오후 2:54

마지막으로 해볼 설릿 정리해봤습니다!

모든 마지막 시도는 0.44나왔던 노트북의 config로 갈겁니다! (epoch, batch size 등등)

1. CLS SEP 토큰을 ##### 으로 통일 + LGBM OPTUNA ON
2. CLS SEP 토큰을 [CLS], [SEP] 으로 지정 (토크나이저 기본 토큰들) + LGBM OPTUNA ON

Text 컬럼이 현재는 Prompt Question + SEP + CLS + Pyspellcheck text의 형식으로 되어 있는데 아래처럼 시도 예정

1. Prompt Question + SEP + CLS + Pyspellcheck text + SEP + CLS + SymSpell text + LGBM OPTUNA ON
2. Prompt Question + SEP + CLS + SymSpell text + LGBM OPTUNA ON
3. Prompt Question + SEP + CLS + SymSpell text + LGBM OPTUNA OFF

혹시라도 다른 방식 시도해보시고 싶은게 있다면 말씀 주세요! (편집됨)

모두 GroupKFold 사용예정입니다



이희상_오프라인 🇰🇷 오후 3:07

일단 아래 3개는 가져가도록 하겠습니다!

답글 2 좋아요



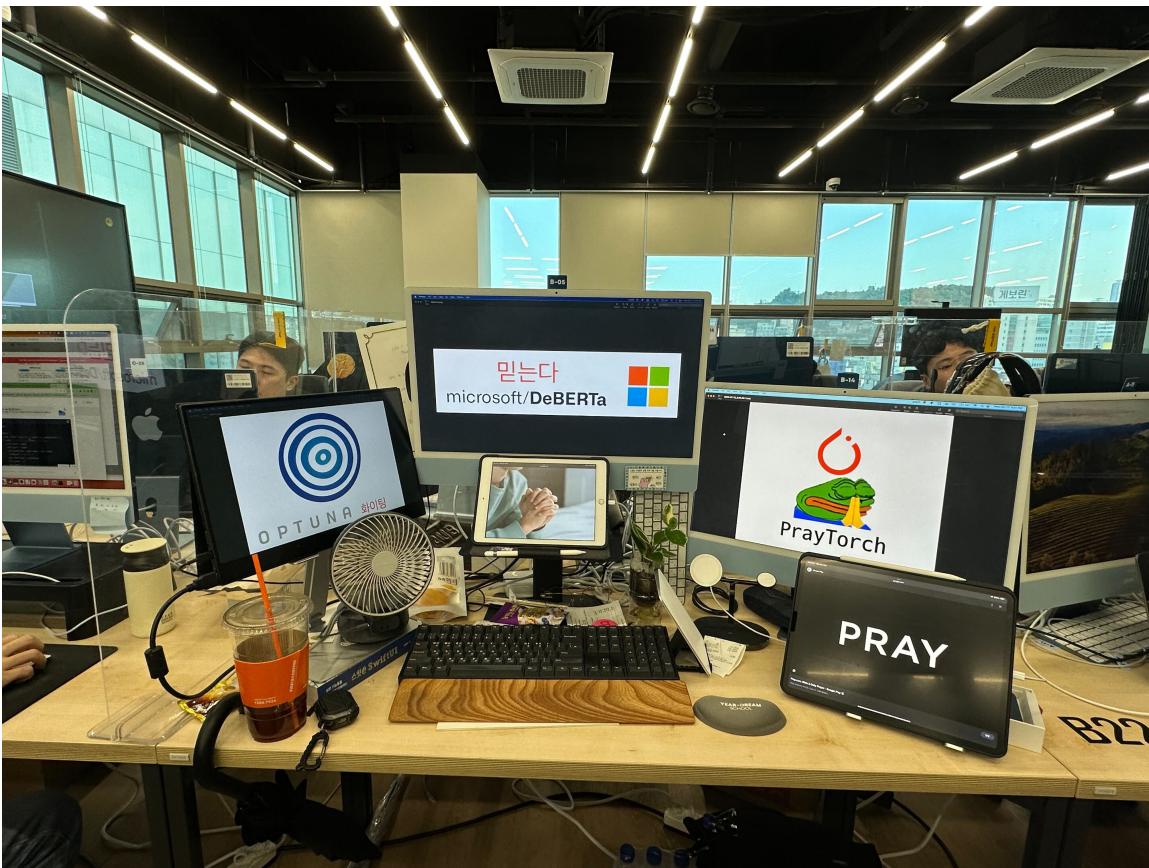
이희상_오프라인 🇰🇷 오후 3:13

지금보니 ##### 과 Group fold 할때가 다 점수가 비교적 괜찮았네요 (그래도 이거라도 알아낸게 있어서 다행...)



이희상_오프라인 🇰🇷 오후 5:05

우리 모두 기도해요



우리에게 남은건 이제 기도밖에 없습니다

임승준_오프라인 오후 5:07
잠들기전에 기도하고 잡니다 ㅋㅋㅋ

한찬혁_오프라인 오후 5:34
진짜, 희상님 한테 많이 배웁니다!!!

이희상_오프라인 오후 5:54
ㅠㅠㅠㅠ

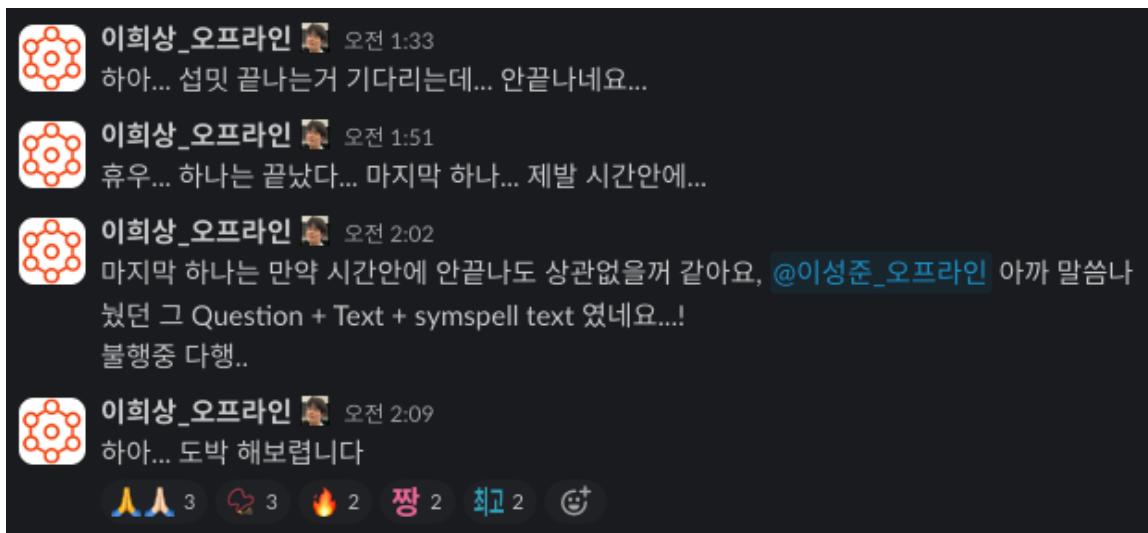
최고 4 ⚡+

이희상_오프라인 오후 9:33
일단 섭밋 예정 노트북 두개 확정입니다

1. 0.44
2. 0.444 - CV 0.46, 0.44에서 오타처리는 Symspell로만 한거입니닷 Optuna X

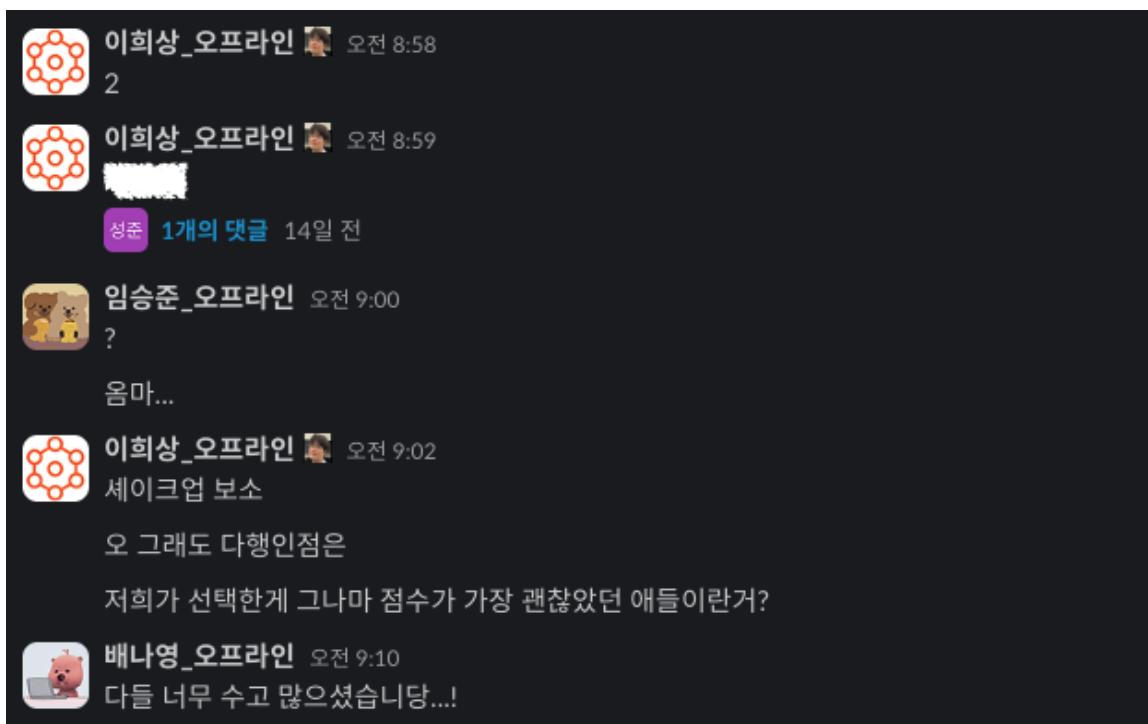
마지막 노트북은 아래 둘 중에 하나...
0.443 - CV 모름, 0.44에서 LGBM Max Depth 만 4 -> 8로 조정
0.445 - CV 모름, 0.44에서 오타처리는 Symspell로만, Optuna 포함
단 현재 public LB 상 0.443까지가 동메달권,
0.443 노트북 대신에 0.445 고르는 건 약간 도박이긴 함... (편집됨)

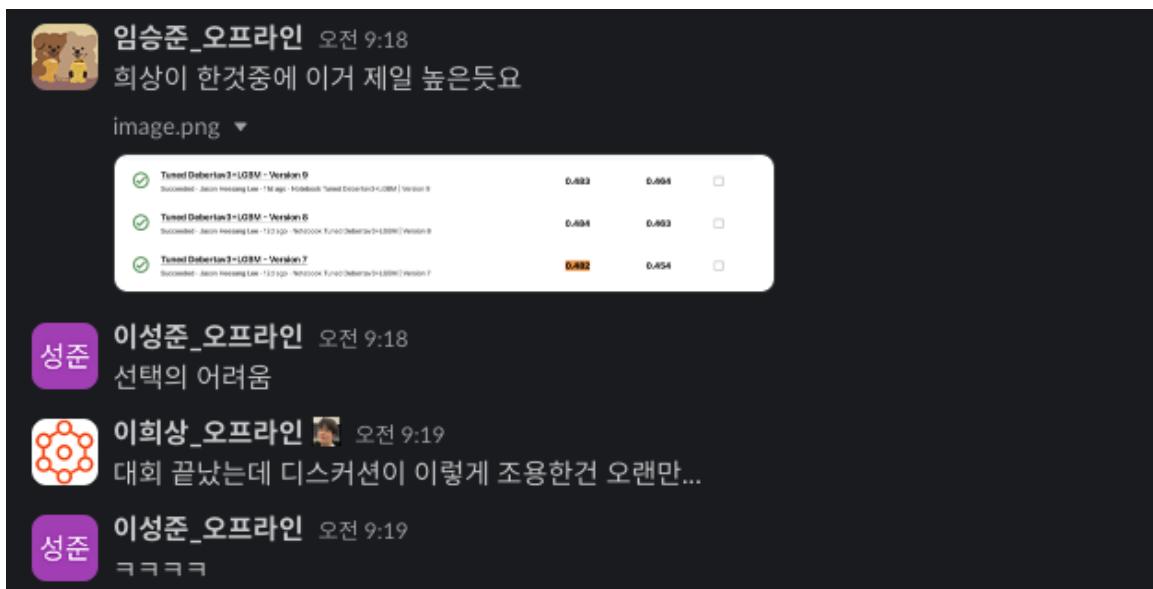
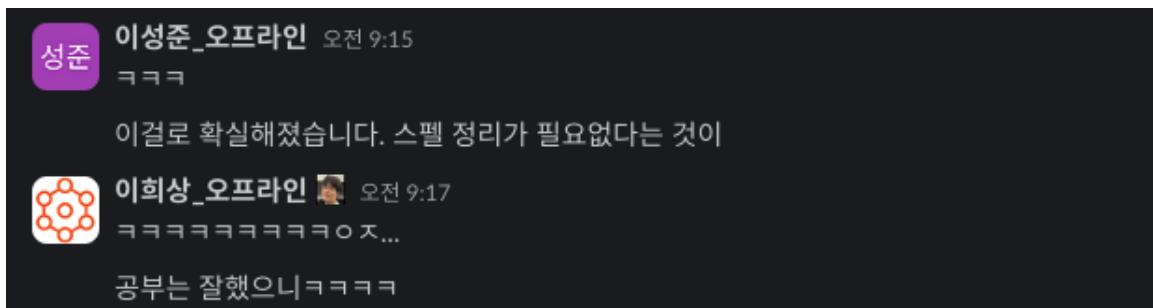
🔥 3 ⚡+



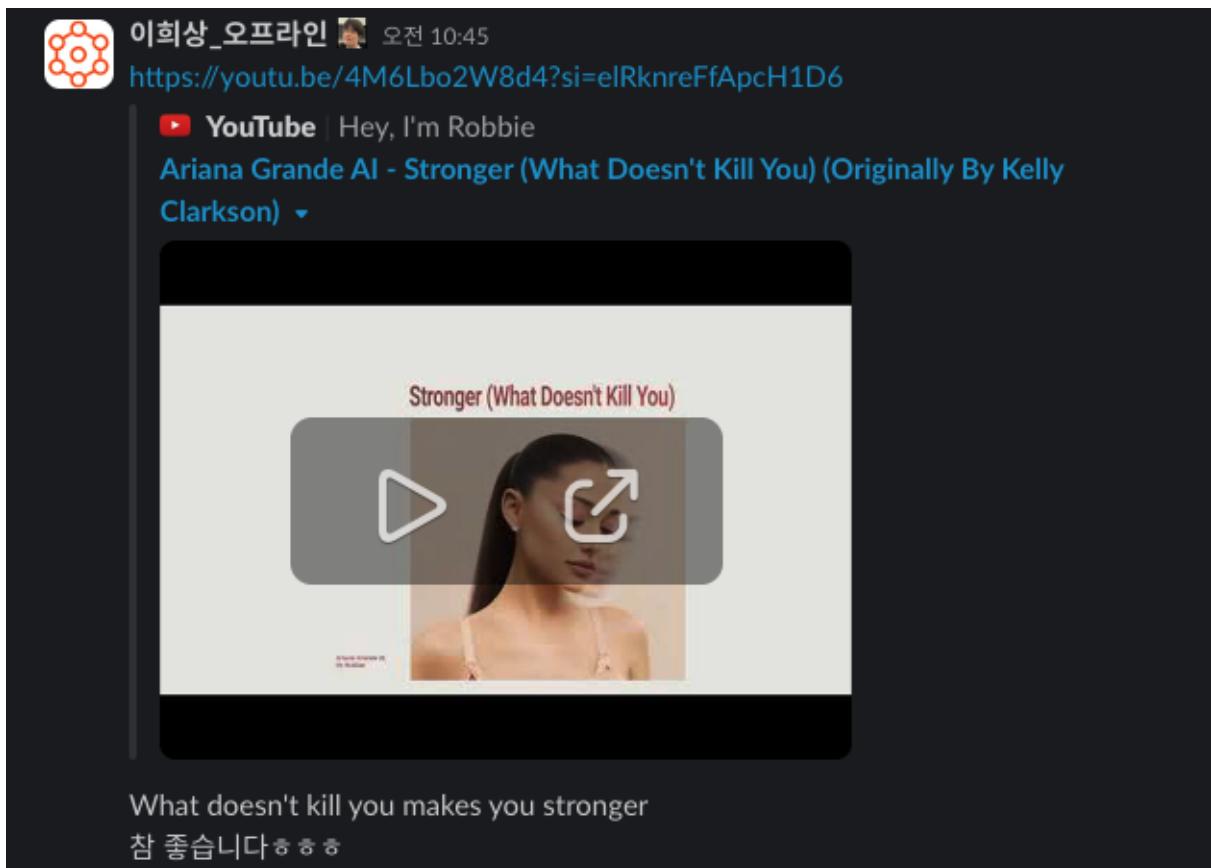
대회 종료 후

▼ 대회가 종료되고 등수가 나온 후... 망연자실 하였지만 그래도 다시 마음을 다잡고 패인 분석을 진행하였습니다.





자세한 패인 분석은 본 페이지 최상단의 PDF 파일 참고 부탁드리겠습니다.



What doesn't kill you makes you stronger

- 나를 죽이지 않는 것은 나를 강하게 만들뿐이다

Stronger의 가사를 되새기며 속상한 마음을 다시 한번 다잡았습니다.