**Wrap Up Report**

**<기술적인 도전>**

LB 점수 0.732, 115등

검증(Validation) 전략

제공된 Dataset에 대해 KFold 5를 적용했습니다. 이것을 적용하기 위해 Dataset Class에서 해당 데이터의 label도 따로 모아 후에 학습할 때 넘겨줬습니다.

사용한 모델 아키텍처 및 하이퍼 파라미터

1. 아키텍처: bert-multilingual-base-cased
2. LB 점수 : 0.7320
3. training time augmentation

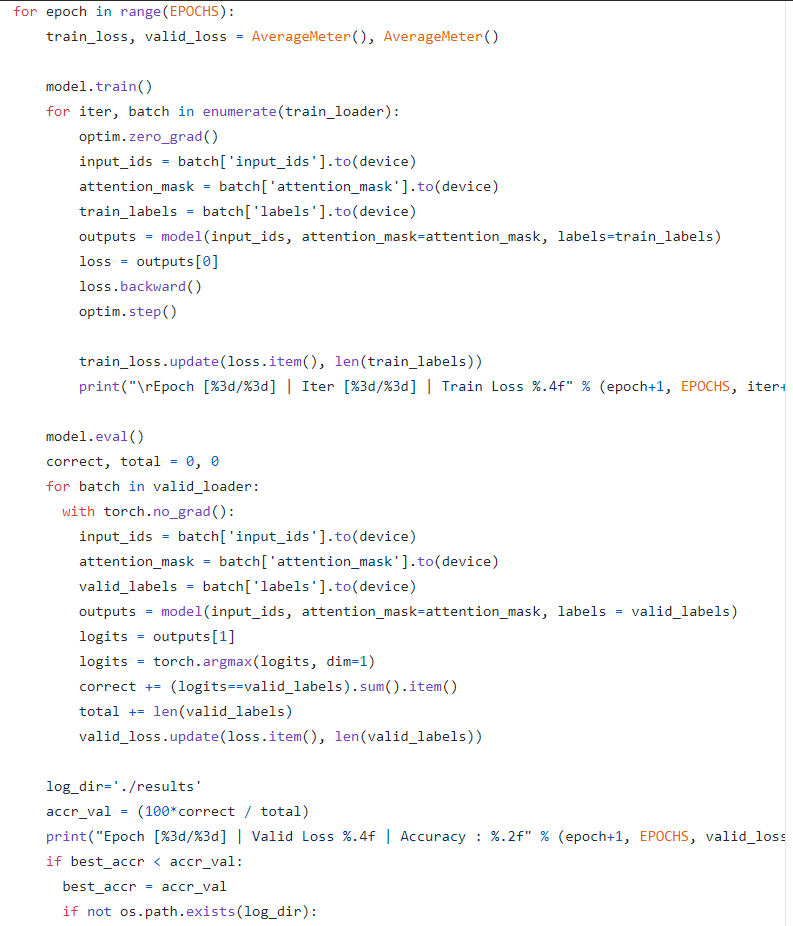
max length : 150

1. 추가 시도

* Batch Size를 처음에는 32으로 시작했다가 16, 64 등 다양한 batch를 시도해 최종적으로 32를 적용해 Local minimum에 빠지는 문제를 해결해 Accuracy를 올렸습니다.
* Learning rate를 1e-5, 5e-05, 5e-06등 다양하게 적용을 해 가장 나은 값이 1e-05라는 것을 알게 되었습니다.
* Optimizer 부분에서 Adam과 AdamW을 두고 여러 번 반복 시도한 결과 AdamW의 결과가 좋다는 결론을 얻었습니다.

1. 직접 Model의 Pipeline을 구성





1. 주어진 Baseline에 있는 TrainArgument와 Trainer를 이용하는 것은 사용에 간편하였지만 제가 원하는 방법을 적용하거나 Hyper Parameter 수정에 어려움을 느꼈습니다.
2. 관련 자료를 보더라도 각 argument의 코드 작동이나 연결에 대한 설명에 이해가 어려워 제가 직접 Train과 Valid에 대한 코드를 구현하기로 했습니다,
3. 모델 구성에 성공해 Train을 적용해 매 checkpoint 마다 model을 저장하는 것이 아닌 Valid에 대한 Best model만을 저장하도록 해 저장 공간 사용을 줄일 수 있고 쉽게 파라미터 수정이 가능해졌습니다.
4. 앙상블(Ensemble) 방법

각 KFold 모델 마다 나온 예측 Label을 K로 나눠 ans에 저장한 뒤 최종적으로 Test가 끝나면 해당 sentence entity relation에 대한 예측 분포에서 Argmax를 이용한 가장 큰 값을 정답으로 사용하였습니다.

1. 시도했으나 잘 되지 않았던 것들
2. 다양한 Hyper Parameter의 수정을 통해 Model의 성능을 올리려 했으나 결국 중요한 건 Data의 다양성과 Base Model의 Task 적합도라는 결론이 나왔습니다.
3. Model을 변환하는 과정에서 중요한 부분의 코드가 변환이 되지 않았는지 기존에 주어진 Baseline Trainer에 비해 같은 Parameter여도 성능이 좋지 않았습니다.