

1.

Viva pro팀 2월 17일 발표를 시작하겠습니다.

2.

우선 목차를 살펴보면 하이퍼파라미터 수정에 관한 내용, 배치사이즈와 f1값의 관계, 말뭉치 태그 보충 관련 내용, 개체명 인식 서비스 개요 순으로 발표하겠습니다.

3.

가장 먼저 수정한 항목은 Epoch입니다. 기본값은 20이며, 이 때 f1 스코어는 0.8763의 값을 가집니다. 30일 때는 0.9119, 35일 때는 0.9112, 40일 때는 0.9120.

조교님이 말씀하신 크게 차이나는 Epoch 값으로 70을 사용하였는데 0.9132의 결과값을 확인하였습니다. 확연하게 다른 f1 값을 기대했었지만, 조금 더 여러가지 값을 테스트 해본 뒤에 loss 까지 고려하여 epoch 값을 선택할 것입니다.

4.

다음으로 수정한 항목은 train batch size입니다. 기본값은 32이며 이때의 f1 score는 0.8763입니다. 처음에는 배치 사이즈를 크게 설정하여 테스트 하고자 했습니다. 그러나 64로 수정한 결과, 0.6371로 크게 떨어진 값이 나타나 오히려 사이즈를 작게 설정하였습니다. 그래서 16일 때, 0.9113이라는 결과값을 얻었고, 작은 값으로 더 다양하게 시도해볼 예정입니다.

5.

다음은 Evaluation batch size입니다. 기본값은 128이며, 64로 줄이거나 256으로 크게 설정해도 f1 score는 0.8763이었습니다. 그러나 loss 까지 고려하면 128 혹은 256 정도의 값을 사용하게 될 것 같습니다.

6.

배치 사이즈를 크게 설정할 때, 왜 F1의 값이 크게 하락하는지에 대한 내용입니다. 해당 부분에 대한 것은 조사한 사실을 토대로 작성하였으나 아직 추측 단계입니다.

우선 ELECTRA는 파인 튜닝 시에 Discriminator 만 사용합니다. 그리고 ELECTRA의 논문에서 언급된 것에 Generator 와 Discriminator 가 가중치를 공유하는 특성이 있습니다.

이러한 가중치에 영향을 미치는 것에는 배치사이즈도 있는데, 배치사이즈는 크기가 작을수록 가중치를 자주 업데이트하고, 클수록 가중치를 덜 업데이트합니다.

결론적으로 배치사이즈를 크게 설정하면 학습 시간은 단축시킬 수 있지만, optimization 과 generalization이 어렵습니다. 그리고 기존에 학습으로 얻게 된 가중치와 배치사이즈를 크

게 설정하여 업데이트가 덜 이뤄진 가중치의 값이 커서 f1 스코어가 제대로 나오지 않는 것으로 예상됩니다.

아마도 기존에 설정된 배치 사이즈의 값이 거의 최댓값에 근접한 것으로 보입니다.

7.

다음으로 부족한 태그에 대한 말뭉치 보충 내용입니다. 기존 문장 수는 131620 문장으로, PT 식물, TR 이론, FD 학문 분야 등에 추가할 문장 수는 15726 문장입니다. 그래서 총 문장 수는 147346 문장입니다.

각 태그 별로 살펴보면 기존 PT 태그 수는 100개 였으며, 이번에 3260개를 추가할 것입니다.

TR 태그는 143개 였으며, 1410개를 추가할 예정입니다.

FD 태그는 394개로 이번에 1175개 추가 예정입니다.

8.

마지막으로 개체명 인식 관련 서비스 내용입니다. 웹페이지에서 시각화 예시로 구성한 화면입니다.

시각화 부분만을 테스트하기 위한 화면이므로 태그가 맞지 않는 점 양해 부탁드립니다.

문장 입력 칸과 결과 창을 나눌 예정이며, 문장을 입력 후에 “클릭” 버튼을 누를 시에 각 개체명들에 레이블링이 되며 색상과 태그 명으로 표시됩니다.

이상으로 발표를 마치도록 하겠습니다.