

프리온 보딩 AI 심화과정 week 1-2

1. Paperswithcode(<https://paperswithcode.com/task/natural-language-understanding>)

)에서 NLU sub task 중 하나를 선택하여 본인 블로그에 정리해보세요. 아래 3가지 항목에 대해서 정리하세요. (각 항목 고려 사항 참고)

1)본인 아이디 및 닉네임

노지훈

2) url

<https://blog.naver.com/nnnobaram/222654592611>

3) 게시물 캡처



1. 문제 정의

Sentiment analysis이란주어진 문장의 극성을 분류하는 과제이다. 예를 들어서 트위터의 트윗을 긍정적, 부정적, 중립적으로 카테고리화 시킬 수 있다. 주어진 텍스트와 라벨을 가지고 정확한 감정을 분석할 수 있다. sentiment analysis의 서브 카테고리에는 다중 감정 분석, 측면 기반 감정 분석, 세분화된 의견 분석, 언어별 감정 분석이 포함됩니다.

2. 데이터 소개(대표적인 데이터 1개)

1) 생성 요약 task를 해결하기 위해 사용할 수 있는 데이터

sentiment analysis를 하기위한 가장 좋은 데이터 셋은 SST 데이터 셋입니다. SST는 **Stanford Sentiment Treebank**의 줄임 말으로서 언어에서 감정의 구성 효과에 대한 완전한 분석을 허용하는 완전히 레이블이 지정된 파스 트리가 있는 코퍼스입니다.

2) 데이터 구조

영화 리뷰에서 추출한 11,855개의 단일 문장으로 구성되어 있습니다. Stanford 파서로 구문 분석되었으며 해당 구문 분석 트리에서 총 215,154개의 고유한 구문이 포함되어 있으며 각 구문에는 3명의 심사위원이 주석을 달았습니다.

각 구문들은 부정, 덜 부정, 보통, 덜 좋음, 좋음 이렇게 나뉘어져있는 sst-5가 있고 또 부정 아니면 긍정인 sst-2가 있다.

3.SOTA 모델 소개(대표적인 모델 1개)

1) task의 SOTA 모델은 무엇인가
smart가 가장 성능이 좋게 나오는 모델이다.

2) 해당 모델 논문이 요약에서 주요 키워드는?


해당 모델의 논문은 SMART: Robust and Efficient Fine-Tuning for Pre-trained Natural Language Models through Principled Regularized Optimization이다.

여기 논문 요약: 전이 학습이 nlp 연구의 판도를 변화시켰다. 그러나 제한된 데이터 리소스와 사전 훈련된 모델의 큰 용량으로 인해서 다운스트림의 데이터를 과적합 시키고 사전 훈련된 모델에 대한 지식을 잃어버리게 한다. 그래서 이를 해결하기 위해서 모델의 용량을 효과적으로 관리하는 Smoothness-inducing regularization, Bregman proximal point optimization로 사전 훈련된 모델에 대해서 잃어버리지 못하게 한다.

여기서 주요 키워드는 사전학습, Smoothness-inducing regularization, bregman 근위점 최적화를 꼽을 수 있겠다.

3.댓글

1) <https://blog.naver.com/codenavy94/222653364998>

**노지훈**


문제 정의에서 개체명 인식에 대한 설명이 예시 까지 들어주어서 이해하는데 도움이 많이 되었습니다.

2022.2.21. 21:20

답글

♡ 0

2) <https://blog.naver.com/codenavy94/222653405430>

**노지훈**


few-shot learning에 대해서 알아갑니다. 도한 논문의 피쳐를 보여줘서 이해를 도와주어서 좋았습니다.

2022.2.22. 20:09

답글

♡ 0

3) <https://blog.naver.com/cardchan/222653567329>

**노지훈**


NLU 에 대한 설명을 시작으로 해서 좋았습니다. 그리고 두가지 모델의 비교값을 보여주어서 이해하기 좋았습니다.

2022.2.22. 19:57

답글

♡ 0

4) <https://blog.naver.com/cardchan/222653484009>

**노지훈**

감성 분석에 정확도와 효율성이 중요해진다는 것을 알 수 있었습니다.

2022.2.22. 20:07

답글

♡ 0