

Assignment 6

영어 디베이트 보조 프로그램: 프로젝트 보고서

영어 분석을 위한 기계학습 (화 123)

202300370 권민서

I. 프로젝트 개요

- 프로그램 종류:

본 프로젝트는 영어 디베이트 수업 준비를 돕기 위한 보조 도구로, 주된 목적은 디베이트 주제에 대해 찬성(pro) 또는 반대(con)의 입장을 자동으로 분류하는 모델을 개발하여, 사용자가 빠르게 디베이트 자료를 분석하고 준비 시간을 단축할 수 있도록 지원함

입력된 문장에 대해 찬성/반대 입장을 자동으로 예측하는 시스템을 구현했으며, 이를 통해 사용자들이 빠르게 반박 포인트를 파악하고 디베이트 준비를 효율적으로 할 수 있도록 돕고자 함

- 해결하려 한 문제:

디베이트 준비는 주로 자료 수집과 분석에서 많은 시간이 소요되며, 입장의 구분 및 반박 포인트 파악이 어려운 경우가 많음. 특히, 디베이트 주제에 대한 여러 입장을 빠르게 분석하고 대응할 수 있는 시스템이 부족한 상황에서, 본 시스템은 입장을 자동으로 구분하여 디베이트 준비 시간을 크게 단축시킬 수 있음.

- 코드 생성 및 수정

프로젝트는 GitHub Copilot 을 이용해 초기 코드를 생성한 후, 모델 학습 및 구현 과정에서 발생한 세부사항 수정과 오류 수정을 직접 수행함.

II. Training

- 데이터 셋

저장 위치: 202300370/assignment4/data-analysis.ipynb

파일명: stance_examples1.csv

구성: sentence, stance, topic

문장 수: 269 개

Stance 분포: pro 와 con 이 균형 있게 분포

토픽 분류: 각 문장은 특정 주제에 대해 찬성(pro) 또는 반대(con)로 분류됨

- 데이터 셋 예시

SENTENCE	STANCE	TOPIC
"WE SHOULD BAN GENE EDITING."	pro	Gene Editing
"GENE EDITING CAN BE EXPLOITED BY THE GOVERNMENT."	pro	Gene Editing
"SPACE EXPLORATION IS MORE IMPORTANT THAN SEA EXPLORATION."	con	Space Exploration

이 데이터 셋을 기반으로 모델을 학습시킴.

- 학습 과정

모델: 사전 학습된 언어 모델(예: DistilBERT)을 기반으로 미세조정(fine-tuning) 진행

입력 포맷: [TOPIC] <topic> [SEP] <sentence>

Optimizer: AdamW

Loss: CrossEntropy

Epochs: 6, Early Stopping 적용

모델 가중치 저장: Google Drive 링크에 저장

저장 위치: 구글 드라이브 링크

모델 학습 과정에서의 주요 파라미터

학습률: 기본 값 (Optimizer 로 AdamW 사용)

배치 크기: 16

에포크 수: 6

Early stopping: 검증 정확도가 3 번 연속적으로 개선되지 않으면 학습 종료

III. Evaluation

- 평가 지표

Accuracy: 48.3%

Macro F1-score: 0.40

Confusion Matrix:

[[115 15]

[124 15]]

- 성능 분석

Confusion Matrix 분석

실제 con(130 개) 중 115 개를 맞추고 15 개는 pro 로 잘못 예측

실제 pro(139 개) 중 124 개를 con 으로 잘못 예측하고 15 개만 맞춤

- 세부 지표

CON 클래스	PRO 클래스
PRECISION: 0.48	Precision: 0.50
RECALL: 0.88	Recall: 0.11
F1-SCORE: 0.62	F1-score: 0.18

- 분석 결과

모델은 con 클래스에는 비교적 정확한 예측을 보였으나, pro 클래스에 대한 예측 성능이 매우 낮았음. 특히 Recall 이 0.11로 매우 낮고, 대부분의 pro 문장을 con 으로 잘못 예측함. 이를 통해 모델이 con 클래스에 치우친 예측을 하고 있음을 확인할 수 있었음.

IV. Inference

- 모델 적용

저장된 모델 가중치를 불러와 새로운 문장에 대해 stance 를 예측한 결과는 다음과 같음.

- 예시 1:

입력: We should ban gene editing. Gene editing can bring ethical problems.

출력: {'stance': 'pro', 'proba_con': 0.468, 'proba_pro': 0.532}

- 예시 2:

입력: Space exploration is more important than sea exploration. We can find useful resources such as minerals or bacteria while exploring the sea.

출력: {'stance': 'con', 'proba_con': 0.509, 'proba_pro': 0.491}

이와 같은 방식으로, 입력된 문장에 대해 pro/con 의 입장을 예측하고 확률 값을 제공하는 시스템을 구현함.

V. 실제 사용

```
... enter the topic: AI should be regulated.  
enter the sources: Prevents job loss, ensures ethical use, avc  
{'stance': 'con', 'proba_con': 0.51, 'proba_pro': 0.49}  
enter the sources: Stifles innovation, regulation lags behind  
{'stance': 'pro', 'proba_con': 0.45, 'proba_pro': 0.55}  
enter the sources: Cyberbullying, comparison culture, addictic  
{'stance': 'pro', 'proba_con': 0.463, 'proba_pro': 0.537}  
enter the sources: AI needs regulation to prevent misuse (deep  
{'stance': 'pro', 'proba_con': 0.494, 'proba_pro': 0.506}  
enter the sources: Arguments against AI regulation often cente  
{'stance': 'con', 'proba_con': 0.589, 'proba_pro': 0.411}  
enter the sources: end
```

```
... enter the topic: we should ban gene editing.  
enter the sources: Curing Diseases: Gene editing could eliminate  
{'stance': 'con', 'proba_con': 0.639, 'proba_pro': 0.361}  
enter the sources: Gene editing can bring ethical problems.  
{'stance': 'pro', 'proba_con': 0.468, 'proba_pro': 0.532}  
enter the sources: Gene editing should be allowed, proponents arg  
{'stance': 'con', 'proba_con': 0.566, 'proba_pro': 0.434}  
enter the sources: Preventing "Genetic Lottery": It could correct  
{'stance': 'con', 'proba_con': 0.578, 'proba_pro': 0.422}  
enter the sources: Gene editing can be exploited by the governmer  
{'stance': 'con', 'proba_con': 0.502, 'proba_pro': 0.498}  
enter the sources: end
```

```
... enter the topic: ocean exploration is more important than space exploration
enter the sources: because the ocean is vital for Earth's climate
{'stance': 'con', 'proba_con': 0.564, 'proba_pro': 0.436}
enter the sources: We can find useful resources such as minerals on the ocean floor
{'stance': 'con', 'proba_con': 0.518, 'proba_pro': 0.482}
enter the sources: ocean exploration would be called more crucial than space exploration
{'stance': 'pro', 'proba_con': 0.495, 'proba_pro': 0.505}
enter the sources: more investment in the ocean will lead to better economic growth
{'stance': 'con', 'proba_con': 0.522, 'proba_pro': 0.478}
enter the sources: Proponents highlight the ocean's role as a source of food and medicine
{'stance': 'con', 'proba_con': 0.544, 'proba_pro': 0.456}
enter the sources: end
```

```
... enter the topic: gun ownership must be restricted.
enter the sources: Arguments for banning guns center on reducing gun violence
{'stance': 'con', 'proba_con': 0.587, 'proba_pro': 0.413}
enter the sources: Public Health Crisis: Gun violence is a major public health problem
{'stance': 'con', 'proba_con': 0.551, 'proba_pro': 0.449}
enter the sources: Many more criminals are killed or wounded each year by guns
{'stance': 'con', 'proba_con': 0.569, 'proba_pro': 0.431}
enter the sources: People also enjoy several of the benefits that gun ownership provides
{'stance': 'con', 'proba_con': 0.568, 'proba_pro': 0.432}
enter the sources: end
```

```
... enter the topic: streaming is better than cable
enter the sources: people would often say that streaming is better than cable
{'stance': 'pro', 'proba_con': 0.402, 'proba_pro': 0.598}
enter the sources: due to its superior flexibility, customization, and lower cost
{'stance': 'con', 'proba_con': 0.518, 'proba_pro': 0.482}
enter the sources: Cable can be better than streaming if you have a large family
{'stance': 'con', 'proba_con': 0.565, 'proba_pro': 0.435}
enter the sources: Reliability: Doesn't depend on your internet connection
{'stance': 'con', 'proba_con': 0.54, 'proba_pro': 0.46}
enter the sources: end
```

VI. 추가 분석

- 성능 개선 필요

모델 성능: Accuracy와 Macro F1-score가 예상보다 낮아, 모델의 성능 향상을 위해 데이터 셋 확장 및 모델 재학습이 필요함.

pro 클래스 Recall: pro 클래스에 대한 예측 성능이 낮아, 데이터 불균형을 해소하고 pro 클래스 문장을 늘리는 작업이 필요함.

- 원인 분석

데이터 불균형: con 클래스 문장이 상대적으로 많아 모델이 con 클래스에 치우쳐 학습됨.

데이터 셋 크기 부족: 269 개의 문장으로는 Transformer 모델을 학습시키기에 부족하여 과적합(overfitting)이나 일반화 실패가 발생했을 가능성이 큼.

입력 포맷 문제: 모델에 토픽 정보가 충분히 반영되지 않았을 수 있음. 토픽을 [TOPIC] <topic> [SEP] <sentence> 형식으로 입력했지만, 이 포맷이 모델에 적절하게 반영되지 않았을 가능성이 있음.

학습 파라미터 최적화 부족: Epoch 수나 Learning rate 등 학습 파라미터가 최적화되지 않아 underfitting 이 발생했을 수 있음.

VII. 개선 방법

- 데이터 확장

더 많은 문장 수집: 뉴스 기사, 논문, 토론 자료 등을 통해 데이터를 확장하고, 특히 con 클래스의 문장을 더 많이 추가해야 함.

데이터 증강: paraphrasing, back-translation 등의 기법을 사용하여 데이터 셋을 더욱 다양화할 수 있음.

- 모델 개선

모델 파인 튜닝: DistilBERT 외에 다른 사전 학습 모델(e.g., BERT, RoBERTa)을 사용하여 성능을 향상시킬 수 있음.

클래스 불균형 처리: 가중치 조정 또는 under 샘플링 / over 샘플링 기법을 사용하여 모델이 더 균형 잡힌 예측을 할 수 있도록 할 수 있음.

학습 파라미터 조정: Epoch 수, 학습률, 배치 크기 등을 Grid Search 나 Random Search 기법을 활용해 최적화할 필요가 있음.

VIII. 결론

본 프로젝트에서는 영어 디베이트 보조 프로그램을 위한 입장 분류 모델을 학습시켜 디베이트 준비 시간을 단축하고, 사용자가 반박 포인트를 빠르게 파악할 수 있도록 지원하는 시스템을 구축함. 모델의 성능은 개선 여지가 있지만, 향후 데이터 확장과 모델 개선을 통해 보다 정확한 예측을 할 수 있도록 지속적으로 개선할 예정.
