

ELECTRA 기반 영화평 감성분석

프로젝트 제안서



<T6>

201611251 공민정

201611276 이규은

201611309 최지현

201612368 이지우

담당교수님 : 김학수 교수님



작품명: *RACCOON*

Review process using

Artificial intelligence tech. to

Classify emotion in

Creative ways for

Opinion mining

Operated by

Neural network





작품 설명

올해(2020), 기존의 문맥 표현 모델들보다 정확도가 높아진 언어 모델 **ELECTRA**(Efficiently Learning an Encoder that Classifies Token Replacements Accurately)가 발표되었습니다.

해당 기술을 적용하여, 이미 존재하는 BERT(Bidirectional Encoder Representations from Transformers)를 사용한 감성 분석 모델을 발전시킨 새로운 모델을 설계해보고자 합니다.

본 작품에서는 영화평 데이터를 바탕으로 긍정/부정적 평가를 구분하는 분류 모델을 개발합니다.

사전 학습한 ELECTRA를 활용한 모델을 사용합니다.

영화평 감성 분석 데이터인 NSMC(Naver Sentiment Movie Corpus)를 훈련 데이터로 사용하며, 성능의 개선을 위해 최신 영화평 데이터를 크롤링해 함께 적용합니다.

기존의 모델보다 높은 성능을 보이는 것이 목표입니다.



새로 만들 SW, 가지고 올 COST SW, HW

새로 만들 SW

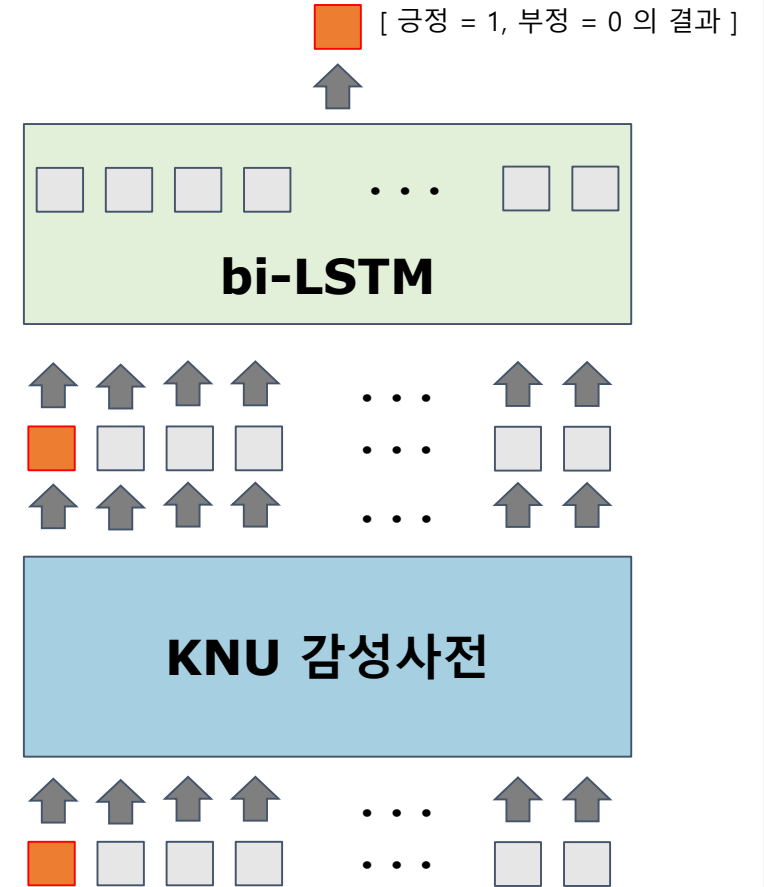
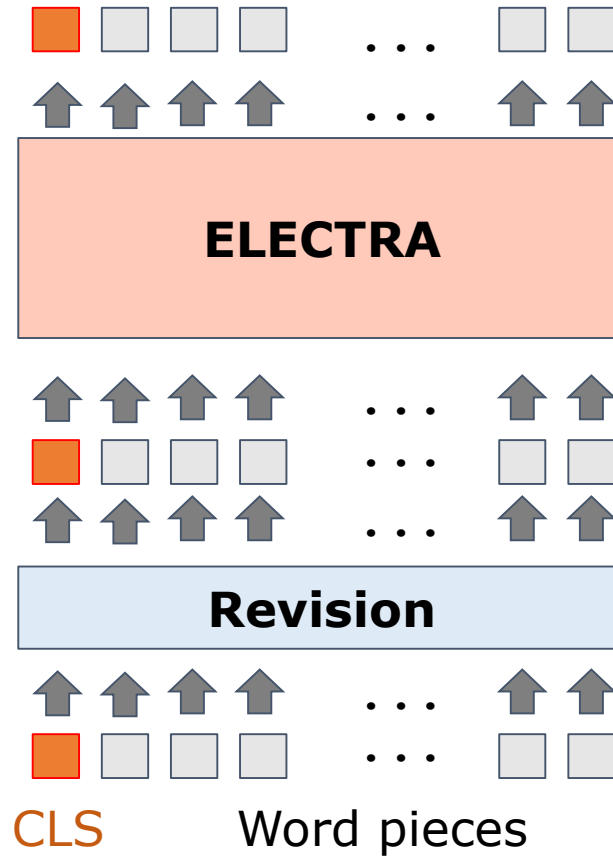
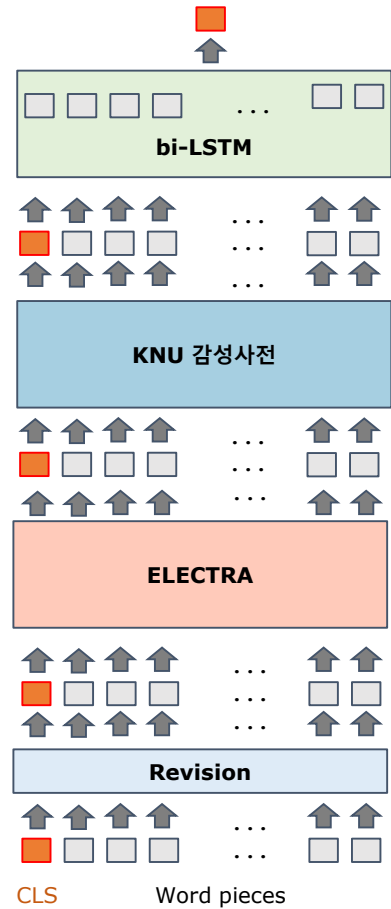
- ELECTRA 기반 Neural Network

가지고 올 COST SW, HW

- AWS
- Pytorch
- KoELECTRA : 한국어 지원 ELECTRA
- NSMC : Naver Sentiment Movie Corpus
- 추가적인 영화평 데이터 (Naver 제외, Test 목적)
- KNU 감성사전 (표준국어대사전을 구성하는 각 단어의 긍부정어를 추출한 사전)



최종 산출물의 형태 및 기능



* CLS : Class. Placed at the beginning of the input example sentence/sentence pair



Alternative Solutions (References, 유사한 것들)

BERT 기반 NN

- Reference 논문

- 1) BERT 기반 Variational Inference와 RNN을 이용한 한국어 영화평 감성분석
- 2) 문맥 표현 기반 한국어 영화평 감성분석



Project Justification

BERT vs. ELECTRA

<Figure 1>을 보면, ELECTRA가 Bert를 포함한 다른 방식들에 비해 매우 빠르게 성능이 향상되며, 동일한 조건 (모델 크기, 데이터, 계산량)으로 비교했을 때 모든 학습 과정에서 더 높은 GLUE(General Language Understanding Evaluation) 성능을 보입니다.

또한, ELECTRA는 RTD(대체 토큰 탐지, Replaced Token Detection) 태스크로 입력의 15%가 아닌 모든 토큰에 대해서 학습하기 때문에 상당히 효율적이면서도 효과적입니다.

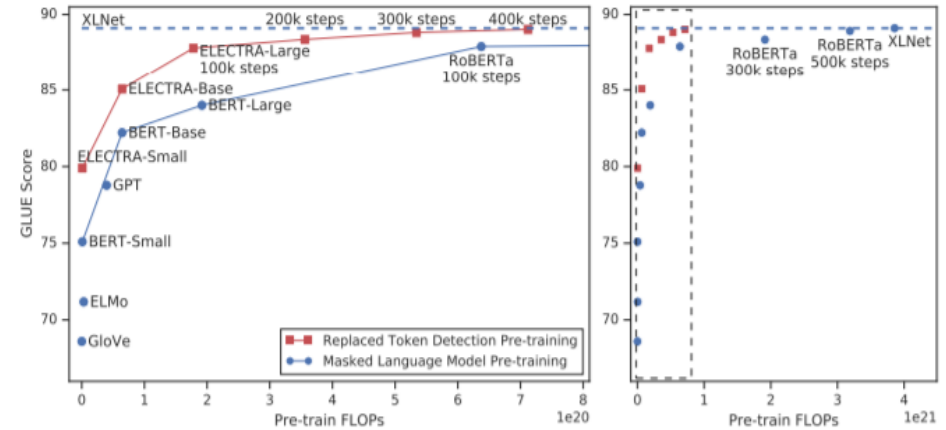


Figure 1: Replaced token detection pre-training consistently outperforms masked language model pre-training given the same compute budget. The left figure is a zoomed-in view of the dashed box.

따라서, BERT보다 본 작품에서의 ELECTRA를 활용하는 모델에서의 분석 결과 더 높은 정확도를 기대할 수 있습니다.



Risk Analysis / Risk Reduction

Risk

1.

Electra는 2020년 3월 발표된 모델이라
자료가 불충분

2.

코로나로 인한 모임 제한으로 소통 불편

3.

딥러닝 추가 공부 필요

Reduction Plan

- 교수님의 조언
- KoELECTRA를 제안한 엔지니어와 연락 시도
- 학습 결과의 높은 정확도를 보장할 수 없으므로 모델을 계속 더 좋은 방향으로 개선 필요

- Notion 활용으로 팀원들간 정보 공유와 소통을 원활하게 함

- Github 상의 오픈소스를 통한 공부
- 방학 중 study 진행



Success Criteria

시간

BERT를 활용한 영화평 감성 분석 모델보다
적은 경과 시간을 기대

정확도

ELECTRA 단일 모델
(정확도: Base - 85.1%, Small - 79.9%)
보다 높은 성능을 기대

Model	Train / Infer FLOPs	Speedup	Params	Train Time + Hardware	GLUE
ELMo	3.3e18 / 2.6e10	19x / 1.2x	96M	14d on 3 GTX 1080 GPUs	71.2
GPT	4.0e19 / 3.0e10	1.6x / 0.97x	117M	25d on 8 P6000 GPUs	78.8
BERT-Small	1.4e18 / 3.7e9	45x / 8x	14M	4d on 1 V100 GPU	75.1
BERT-Base	6.4e19 / 2.9e10	1x / 1x	110M	4d on 16 TPUv3s	82.2
ELECTRA-Small	1.4e18 / 3.7e9	45x / 8x	14M	4d on 1 V100 GPU	79.9
50% trained	7.1e17 / 3.7e9	90x / 8x	14M	2d on 1 V100 GPU	79.0
25% trained	3.6e17 / 3.7e9	181x / 8x	14M	1d on 1 V100 GPU	77.7
12.5% trained	1.8e17 / 3.7e9	361x / 8x	14M	12h on 1 V100 GPU	76.0
6.25% trained	8.9e16 / 3.7e9	722x / 8x	14M	6h on 1 V100 GPU	74.1
ELECTRA-Base	6.4e19 / 2.9e10	1x / 1x	110M	4d on 16 TPUv3s	85.1

Table 1: Comparison of small models on the GLUE dev set. BERT-Small/Base are our implementation and use the same hyperparameters as ELECTRA-Small/Base. Infer FLOPs assumes single length-128 input. Training times should be taken with a grain of salt as they are for different hardware and with sometimes un-optimized code. ELECTRA performs well even when trained on a single GPU, scoring 5 GLUE points higher than a comparable BERT model and even outscoring the much larger GPT model.



References

- ELECTRA: <https://openreview.net/forum?id=r1xMH1BtvB>
- Character-Level Korean ELECTRA Model: <https://github.com/monologg/KoCharELECTRA>
- 박천음, 이창기, "BERT 기반 Variational Inference와 RNN을 이용한 한국어 영화평 감성 분석", 2019
- Kevin Clark, Minh-Thang Luong, Quoc V. Le, Christopher D. Manning, "ELECTRA: Pre-training Text Encoders as Discriminations Rather Than Generators", 2020