

딥러닝을 이용한 배달 음식점 리뷰 자동 생성

데이터 셋

- 요기요 웹 페이지 치킨 음식점 리뷰 데이터 -> 총 126,800개

전처리

- 중복 데이터와 치킨 음식점 리뷰를 제외한 나머지 리뷰들을 제거
- 100자 이상 긴 문장을 넣을 시 과적합이 발생하여 글자 수 25자 ~ 100자 사이의 리뷰를 사용

모델

- Seq2Seq를 이용한 LSTM,
- KoGPT-2

영화 리뷰 감성분석을 위한 텍스트 마이닝 기반 감성 분류기 구축

데이터 셋

- 최근 2주 영화 리뷰 데이터를 크롤링하여 125만개의 데이터를 수집
- 모델 학습을 위해 1만개를 사용(9000개는 train, 1000개는 validation)

전처리

- 토큰화

모델

영화 리뷰 감성분석을 위한 텍스트 마이닝 기반 감성 분류기 구축

데이터 셋

- 최근 2주 영화 리뷰 데이터를 크롤링하여 125만개의 데이터를 수집
- 모델 학습을 위해 1만개를 사용(9000개는 train, 1000개는 validation)

전처리

- 토큰화

모델

감성 분류를 위한 모델 → BERT

문장 생성을 위한 모델 → BERT, GPT