

Delete-MASS Gen : MASS를 이용한 단어 n-gram 삭제 및 생성 기반 한국어 스타일 변환

최형준¹, 나승훈²
12 전북대학교

modera2017@gmail.com, nash@jbnu.ac.kr

I. 스타일 변환

긍정적인 문장을 부정적으로 바꾸는 등의, 문장이 가진 속성을 변화시켜 새로운 문장을 생성하는 작업.

부정 -> 긍정	
쓰레기 수준은 아니라도 참 돈 아깝게 본 영화다.	정말 감동은 아니라도 재밌게 보고 본 영화다.
보는 내내 힘들었다...	보는 내내 재밌었다...
긍정 -> 부정	
내용도 있고 좋군요	재미없다. 쓰레기 영화
또 하나의 교훈을 준 영화	1점도 아까운 쓰레기 영화

II. 감성분석기

문장이나 각 토큰이 해당 스타일을 얼마나 가지고 있는지 측정하기 위해 감성분석기를 사용. 감성분석기가 문장을 각 스타일이라고 예측한 확률을 그 문장이 해당 스타일 속성을 가지고 있는 정도로 사용

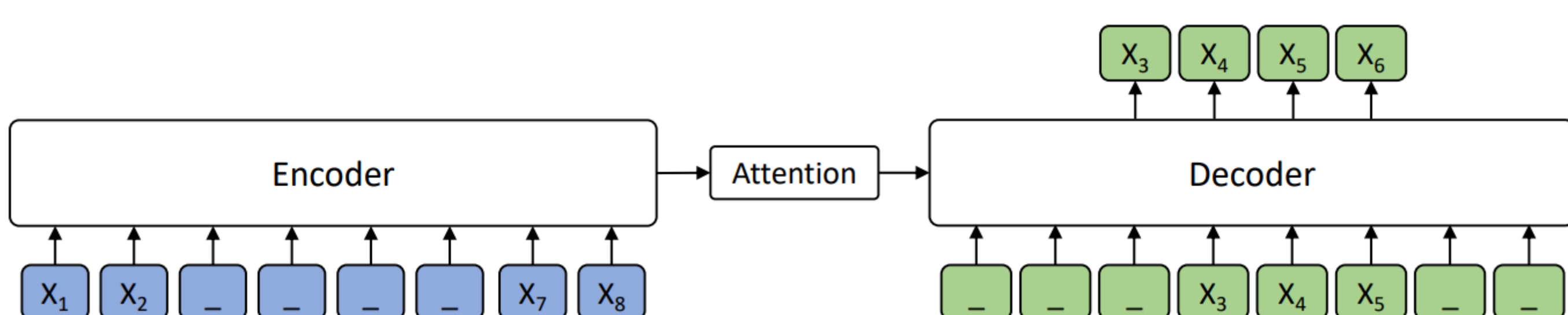
문장	긍정	부정
깔끔하고 세련된 유머와 복잡하지 않은 인물전개 굿!	0.9976	0.0024
스릴러의 모범답안 이런 영화가 국가대표 보다 평점이 낮다니..	0.9526	0.0474
최악의 시작에서 최악의 마무리로	0.007	0.9993
삼류도 아니고 계속 불쾌해지는 영화네	0.0041	0.9959

단어 n-gram에 대해 해당 n-gram이 가진 스타일 속성은 그 n-gram을 삭제한 문장을 감성분석기가 평가했을 때 감소한 확률만큼 존재한다고 보고 각 n-gram별 점수를 측정

n-gram	긍정	부정
최 고	0.9916	0.0084
재 미 있	0.8972	0.1028
종 은 영화	0.8536	0.1464
최악 의	0.0008	0.9992
기 대 이하	0.2742	0.7253
줄 았	0.1504	0.8496

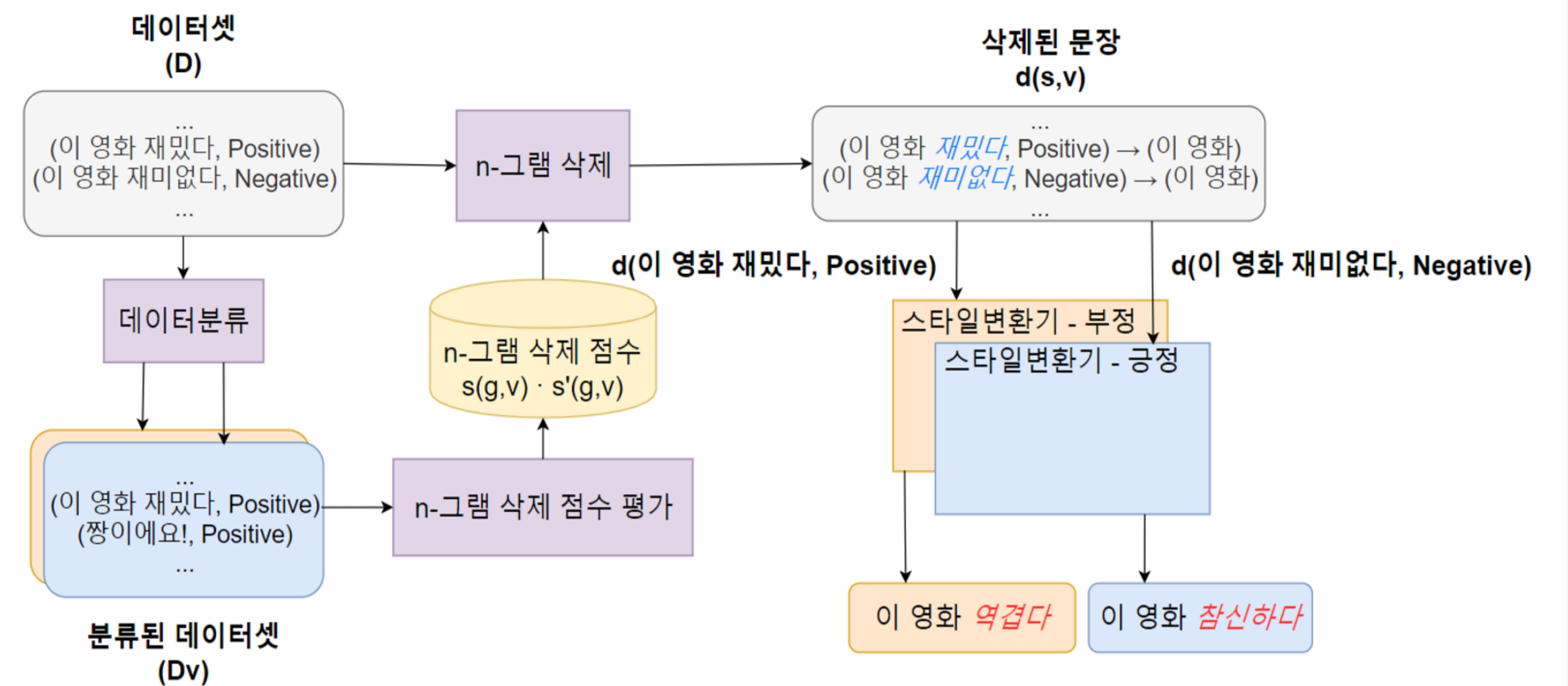
III. 모델

-MASS Masked ML을 통해 Seq2Seq 모델을 선행학습 시켜둔 것. Masked ML은 문장의 무작위 한 부분을 [MASK] 토큰이나 전체 단어 집합 중 무작위 단어 또는 원래 단어로 마스킹 한 후, 마스킹된 부분을 생성시키도록 학습. 선행 학습된 Seq2Seq 모델을 fine-tuning할 경우, 기존의 모델에 비해 번역 등의 문장 생성 성능이 더 뛰어나



스타일 변환을 위해 다음 3가지 작업을 시행

- 삭제 문장으로부터 스타일 변환할 스타일을 가진 n-gram을 삭제
- 복원 삭제된 문장을 원래 문장으로 복원시키도록 모델 학습
- 변환 삭제된 문장을 변환할 스타일의 모델로 변환된 문장 생성



IV. 실험

- 데이터셋 네이버 한국어 영화 리뷰 데이터셋을 사용 각 영화 리뷰에 대해 긍정/부정 레이블을 포함 각 영화리뷰는 한국어 형태소 분석기를 통해 토큰화 후 BPE를 통해 한번 더 토큰화 시 행 BPE 처리 후 길이 5 미만인 데이터는 폐기
- 모델 하나의 트랜스포머를 통해 구현된 MASS로 구성 긍정/부정 각각 하나의 개별적인 모델을 둠

■ 삭제범위 별 스타일 변환 성능 - 감성분석기 출력

스타일	평균점수	평균변화값
긍정 원문	0.192415	-
부정 원문	0.200523	-
긍정 → 부정 MASS 30만개	0.848219	0.655804
긍정 → 부정 MASS 50만개	0.906090	0.713675
부정 → 긍정 MASS 30만개	0.838447	0.637924
부정 → 긍정 MASS 50만개	0.817979	0.617456
긍정 → 부정 4만개 [1]	0.717528	0.569260
부정 → 긍정 4만개 [1]	0.722657	0.556466
긍정 30만개 삭제 후	0.474009	0.281594
긍정 50만개 삭제 후	0.470049	0.277634

■ 감성분석기 출력의 변화값 분포

