**모델링 결과 보고서**

### 

## 1.평가 데이터 셋



‘무신사’ 플랫폼에서 무신사 스탠다드,국내 SPA 세 브랜드의 상의,바지,아우터 카테고리의 판매 순위 상위 10개의 상품에서 약 2만개의 리뷰 데이터를 추출하고,

그중 랜덤으로 2000개의 데이터를 뽑아 평가 진행

**데이터의 레이블**

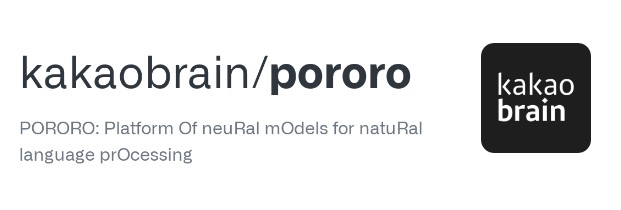
| Positive | Negative |
| --- | --- |
| 1969 | 31 |

## 2.평가 지표

정확도(accuracy)

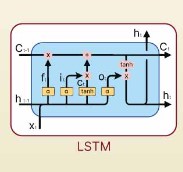
리뷰데이터의 감정 분류 TASK로 긍정부정으로 나누기 떄문에 정확도를 지표로 설계

## 3.사용한 모델



**pororo**

카카오브레인에서 개발한 자연어 프레임워크, 오픈소스



**LSTM** word2Vec, mecab

자연어처리에 적합한 시퀀스 데이터 처리에 효과적인 모델



**BERT**(Bidirectional Encoder Representations from Transformers)

구글이 공개한 사전 훈련된 모델, TASK에 맞게 파인튜닝이 가능한 모델

세가지 모델 모두 오픈 데이터로 공개되어있는 네이버 쇼핑 리뷰 데이터셋으로 학습을 진행, 긍정과 부정으로 나누어진 약 20만개의 데이터로, 진행하는 프로젝트와 특성이 비슷한 특징이 있어 학습진행함.

|  | 별점 | 건수 |
| --- | --- | --- |
| 긍정 (99,963) | 5 | 81,177 |
| 4 | 18,786 |
| 부정 (100,037) | 2 | 63,989 |
| 1 | 36,048 |
| 계 | | 200,000 |

## 

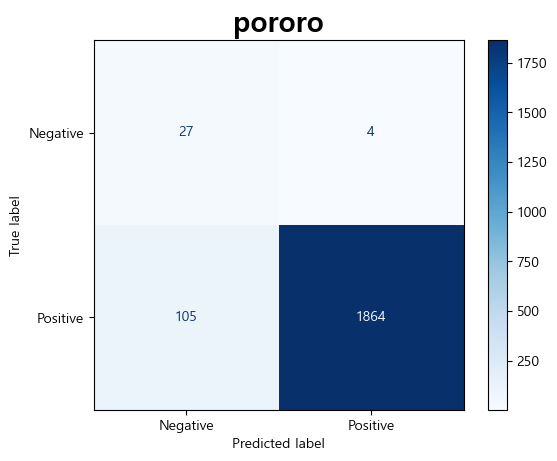
## 4.평가표(정확도,Confusion matrix)

**HUMAN(기준)**

| Positive | 1969 |
| --- | --- |
| Negative | 31 |

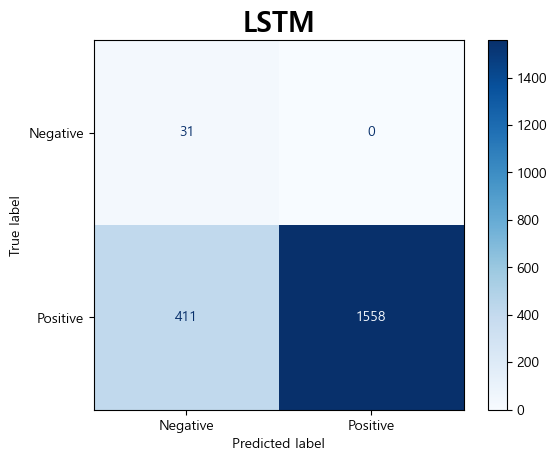
**PORORO(accuracy:94%)**

| Positive | 1868 |
| --- | --- |
| Negative | 132 |



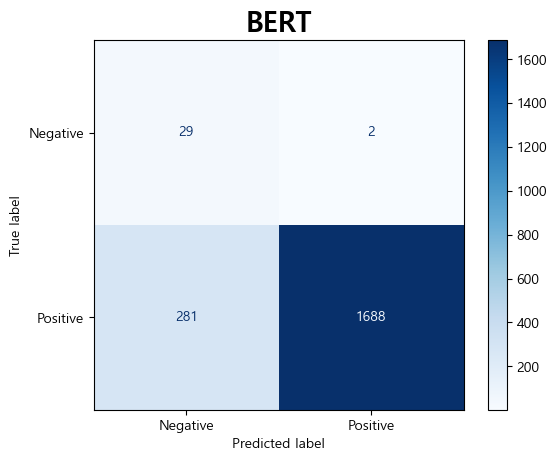
**LSTM(accuracy:79%)**

| Positive | 1690 |
| --- | --- |
| Negative | 310 |



**BERT(accuracy: 85%)**

| Positive | 1558 |
| --- | --- |
| Negative | 442 |



2000개의 데이터로 테스트한 결과 pororo의 정확도가 제일 높은 것을 확인 할 수 있다

## 5.추후 개선사항

데이터 셋 레이블의 편중이 Positive에 치우쳐 정확하게 분류하는지 알기 어렵다.

정확도가 높은 모델들이지만 추후 F1 SCORE로 평가해보면 좋을 것 같다.