

# 일 일 보고서

프로젝트	무신사 온라인 소비자 리뷰 기반 데이터 분석		
작 성 자	team3	일 시	2025-11-05

금 일 실 시 사 항	비 고
<p>[곽동원]</p> <ul style="list-style-type: none"><li>- 데이터 정제 규칙 기반 전처리 코드 확정 및 공유</li><li>- BERT 모델에 파인튜닝 할 데이터 Kc-BERT 로 분류</li><li>- BERT 모델 학습 및 BERT 모델 사전학습 실시</li><li>- ELECTRA 모델 학습</li></ul> <p>[강지연]</p> <ul style="list-style-type: none"><li>- 데이터 정제 규칙 기반 전처리 코드 확정 및 공유</li><li>- BERT 모델에 파인튜닝 할 데이터 Kc-BERT 로 분류</li><li>- BERT 모델 학습 및 BERT 모델 사전학습 실시</li><li>- ELECTRA 모델 학습</li></ul> <p>[안호용]</p> <ul style="list-style-type: none"><li>- 데이터 정제 규칙 기반 전처리 코드 확정 및 공유</li><li>- BERT 모델에 파인튜닝 할 데이터 Kc-BERT 로 분류</li><li>- BERT 모델 학습 및 BERT 모델 사전학습 실시</li><li>- ELECTRA 모델 학습</li></ul> <p>[정수아]</p> <ul style="list-style-type: none"><li>- 데이터 정제 규칙 기반 전처리 코드 확정 및 공유</li><li>- BERT 모델에 파인튜닝 할 데이터 Kc-BERT 로 분류</li><li>- BERT 모델 학습 및 BERT 모델 사전학습 실시</li><li>- ELECTRA 모델 학습</li></ul>	

<p><b>상세 실시 사항</b></p> <ol style="list-style-type: none"> <li>1. 데이터 정제 규칙 기반 전처리 코드 확정 및 공유             <ol style="list-style-type: none"> <li>1.1 사전학습 데이터셋 : 구어체, 뉴스 원문 및 뉴스 리뷰, 영화 및 게임 리뷰                 <ul style="list-style-type: none"> <li>-&gt; Pre-training_dataset.txt</li> </ul> </li> <li>1.2 파인튜닝 데이터셋 : 네이버 쇼핑 리뷰                 <ul style="list-style-type: none"> <li>-&gt; Fine-tuning_shopping.txt</li> </ul> </li> </ol> </li> <li>2. BERT 모델에 파인튜닝 할 데이터 Kc-BERT 로 분류             <ol style="list-style-type: none"> <li>2.1 Kc-BERT 제작자 Git-hub 참고</li> <li>2.2 “KcBERT: 한국어 댓글로 학습한 BERT” 논문 참고</li> <li>2.3 파인튜닝 할 데이터 Kc-BERT 로 분류</li> </ol> </li> <li>3. BERT 모델 학습 및 BERT 모델 사전 학습 실시             <ol style="list-style-type: none"> <li>3.1 Kc-bert 제작자 Git-hub 사전학습 방법 참고</li> <li>3.2 사전학습 test 진행</li> </ol> </li> </ol>	
명 일 예 정 사 항	비 고
<ul style="list-style-type: none"> <li>● BERT 사전 학습 실시</li> <li>● 파인튜닝 방법 학습</li> <li>● ELECTRA 모델 학습, 논문 리뷰</li> <li>● 크롤링 데이터 재정의</li> </ul>	

이 슈 사 항	