|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 프로젝트 개발 일지 | | | | |
| 프로젝트명 | 증권뉴스의 감성분석을 통한 주가 예측 프로젝트 | | | |
| 대분류 | 데이터 분석, 감성 분석 | 문서번호 |  | |
| 모듈명 |  | 작성자 | 유청환 | |
| 작성일 | 2022-10-13 | 문서책임자 |  | |
|  | 개 발 내 용 | | | 비 고 |
| 당일 내용 | 1. 전처리 완료된 데이터로 머신러닝 모델 제작 2. 형태소 분석, 토큰화, 불용어 제거 등의 데이터 정제과정 진행 3. TF-IDF로 중요 단어 통계분석 및 로지스틱 회귀를 사용해 모델 제작 및 저장(64%의 정확도를 가진 모델 제작 완료) 4. 제작된 모델을 사용하는 함수 제작 | | |  |
| 익일 계획 | 데이터 분석(감성분석, 키워드 분석) | | |  |
| 산출물 | 1. **토큰화, 불용어 제거 등의 데이터 정제과정 진행**    1. 정제과정 진행 후 토큰 개수 그래프(가장많이 나온 50개의 토큰)      * 1. 정제과정 진행후 토큰 개수 시각화(가장많이 나온 50개의 토큰)      1. **형태소 분석 진행 및 머신러닝**   2-1. 코드    2-2. 과정   * 수집한 데이터를 학습데이터셋, 테스트 데이터셋으로 8:2 비율로 분류해 머신러닝 진행 * 뉴스타이틀 별로 나오는 단어를 파악해서 수치화(TF-IDF 사용), 학습진행(로지스틱 회귀 사용) * 모델 저장   2-3. 결과   * 약 64%의 정확도를 가진 모델제작 완료  1. **모델 사용**   3-1. 코드    3-2. 문제점   * 직접 실행해본 결과 100% 부정의 결과가 나옴 * 다른 뉴스 데이터를 수집하고 모델에 적용 시켜 코드의 정확도를 확인하는 작업 필요 | | | |
| 참고 사항 | 데이터 분석 및 모델 제작 완료 | | | |