**주간 업무 보고서**

| 프로젝트 | **위니텍 자연어 처리 기반 유사 항목 도출 모델 분석 프로젝트** | | |
| --- | --- | --- | --- |
| 기간 | 2023년 05월 02 일 ~ 2023년 05월 13일 | 작성자 | 임 도 영 |
| 업무 내용 | | 완료 여부 | 비고 |
| 금주  진행  상황 | 프로젝트 상세기능 정의 및 모델 관련 공부 | 진행중 | 자료수집 시 기업에서 필요한 주제의 3단 법 자료 몹시 부족. |
| 국가법령정보센터에서 데이터 수집 크롤링 하여 csv 변환   * Selenuim 사용 * 수집기간 : ‘23.05.04 ~ 23.05.11 * 데이터 건수: 법령 45건   (임도영, 정도, 김도형) | 진행완료 |
| csv 전처리 (법 조문과 내용 분리)   * 판다스 사용   (임도영, 정도, 김도형) | 진행완료 |
| 기획서 작성 (임도영) | 진행완료 |
|  |  |
|  |  |
|  |  |
|  |  |
| 차주  계획 | 데이터 전처리 예정   * 결측치 제거, 삭제된 법령 제거, 데이터 중복 처리, 공백제거, 토크나이징, 자연어 등   (김도형) | 진행 중 |  |
| 유사도분석 코드 작성   * 자카드 분석, 코사인 분석   (정도) | 진행 중 |
| 단어, 문장 임베딩  (임도영) | 진행 예정 |
|  |  |
|  |  |
| 지도  교수  피드백 | 위니텍 멘토링 피드백 못받음 |  |  |

**[금주 진행 사항 세부 내용]**

작성 내용

* 금주 진행 중 자료 조사 내용: 웹사이트 링크 및 내용 등을 간단히 작성
* 프로젝트 계획서 상에 정의한 개인별 업무 분담 관련 진행 내용
* 프로젝트 진행 도중 필요한 개인별 조사 내용 등 자유 형식으로 작성하면 됨
* 노션 링크 ㄱㄱ