**회　의　록**

|  |  |
| --- | --- |
| 회　의　명 | 산학 연계 SW프로젝트 int조 회의 및 교수님들과 미팅 |
| 일　　　시 | 2021.02.24 16:30~17:30 |
| 장　　　소 | 8퍼센트 회사 건물 내 |
| 참　석　자 | 한정수, 권나현, 윤서안, 조민경 |

|  |  |
| --- | --- |
| 회의 내용 | 1. 서론  - 프로젝트 주제와 제목은 바꿔도 상관없음.  - SNS만을 이용하기 때문에 직종이 제한적임.  - 어떤 factor들이 가치 있는지가 중요함.  - 회사가 이 프로젝트를 통해 인사이트를 얻을 수 있는 정도(프로토타입 수준)를 바람.  2. 주제 관련 부분  - 신용도 관점에서 가장 좋은 것은 변화가 적은 것임. 예를 들어 수입이 늘어나는 것도 좋지만 무엇보다 더 줄어들지 않고 유지하는 것이 중요함.  - 과거와 최근을 비교하는 것(변화량)을 중점적으로 봐야함. 예를 들어 포스팅 빈도나 댓글과 좋아요 수 증감 등이 있음.  - 성실도와 지속 가능성의 경우 인스타그램 활동의 active함(이벤트성 게시글과 같은 평소와 다른 이미지의 게시글 업로드 등)과 꾸준함으로 알 수 있음.  3. 기술적인 부분  - 크롤링으로 할 수 없는 부분까지 API로 간단하게 수집 가능  - 주피터 노트북에서 selenium의 headless를 사용하면 웹 브라우저 창을 띄우지 않고 프로그램 실행이 가능  - Pandas 라이브러리를 사용하여 데이터베이스 없이 데이터프레임으로 바로 분석이 가능  - 일정 기간 동안의 데이터를 축약해서 sum, avg 등 미리 만들어 놓으면 훨씬 덜 복잡하게 분석 가능 |

|  |  |
| --- | --- |
| 회의 내용 | 4. 결과물  - 회사 측에서 크게 두가지의 결과물을 예시로 언급  (1) 전체 데이터(인스타그램 계정들)의 주요 factor들 분석  5개 정도의 주요 factor들을 뽑고 factor들의 가치와 이들에 대한 데이터의 분포를 확인할 수 있는 결과물  (2) 하나의 계정을 선택했을 때의 데이터 분석  하나의 계정이 선택되면 그에 따른 (1)에서의 주요 factor들과 계정의 상대적 위치를 확인할 수 있는 결과물(row data도 있으면 좋음) |
| 향후 일정 | 2021.02.26 오후 1시에 조원들끼리 온라인 회의 예정 |