**[SSAFY] 특화 프로젝트 멘토링 활동일지**

**작성일자: 2022 년 3 월 10 일**

|  |  |  |  |
| --- | --- | --- | --- |
| **도메인** | **빅데이터 추천** | | |
| **멘토** | **서동완** | **팀 코드** | **A306** |
| **진행 일시** | 2022년 3월 10일 | | |
| **진행 내용** | **📌 전문가 리뷰 및 Q&A**  빅데이터 구현은 다음 순서를 따른다.  **1. 데이터 필드 정의**   * **어떤 데이터필드를 사용해서 게임을 정의할 것인가** * 데이터에 따른 결과를 다 뽑아놓은 상태여야 한다   + 별점 → 5종류     - 결과값이 이미 나와있는 상태     - 다른 데이터를 넣는다고 해도 종류가 한정적임   + 사람이 만명이어도 결과는 수천가지 이내 → 미리 만들어둔 출력값(정적)에서 도출   **2. 알고리즘**   * 알고리즘끼리 섞어서 쓸 수도 있음 * 알려진 알고리즘은 오픈소스를 사용하는게 좋다 * 1에서 정의한 데이터 필드를 가지고 협업필터링이 가능한지는 직접 파악이 필요   **3. 구현 방법**   * python은 결과만 만들어서 spring boot에 올리는 방식으로 구현 가능   1. 수작업   2. 파일 공유   3. 코드 or 커맨드 한두 줄 * 장고 서버를 따로 띄울 필요 없다   1. 투자하는 노력 및 시간 대비 메리트가 적음   2. 타임 딜레이 발생 * 실시간 분석의 필요성?   1. db화를 시키면 편하긴 하지만   2. 파일로 읽어서 해도 됨      + 엑셀 → 읽어서 결과 도출 → 저장 → 웹에서 dp   **+. 빅데이터 분석에서 필요한 고민**   1. 빅데이터의 결과를 보고 이것이 사용 가능한지, 왜 괜찮은지에 대하여 2. 오픈소스의 input data와 우리가 쓰는 input data가 얼마나 잘 매칭될 수 있는가    1. 이에 대한 설명 3. 사용자에게 유의미한 결과를 어떻게 dp할 것인가 4. 얼마나 유의미한 분석을 했는가 5. 우리가 만든 알고리즘이 얼마나 그 개념에 부합하는가 6. 이 결과가 얼마나 유의미한가 | | |
| **특이 사항** | 다음 리뷰때는 해온 것들을 보여줄 것   * 알고리즘의 수식적 정의 * 알고리즘 설명 링크 * 데이터필드 정의 등 | | |