|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **최종 보고서** | | | | | | | | |
| 프로젝트 명 | 국문 | | | 쿠팡 상품평 분석 프로젝트 | | | | |
| 영문 | | |  | | | | |
| 수행기간 | 2022 . 09 .12 ~ 2022. 10. 3 | | | | | | | |
| 참여자 | 소속 | | | | 참여자성명 | 휴대전화 | 이메일 |
| 4조 | | | | 유청환 | 010-4158-1389 | katimere15@gmail.com |
| 4조 | | | | 정희 |  | Heejung130a@gmail.com |
| **개발결과요약** | |  | | | | | | |
|  | |  | | | | | | |
| 1. 개발 결과물에 대한 요약    1. 입력 화면      * 1. 메인 화면      * 1. Analysis\_box      * 1. 데이터 흐름     2. 요구사항 분석 대비 핵심 개발 사항   |  |  |  | | --- | --- | --- | | 순번 | 계획서 상 요구사항분석 내용 | 요구사항 대비 개발 사항 | | 1 | 계획서 제작 및 설계 | 담당 : 팀 전체  Google presentation으로 진행 | | 2 | 데이터 수집 및 데이터 베이스 | 담당 : 팀 전체  mysql 사용   1. database model | | 3 | 상품평 분석 | 담당 : 유청환  KeyBERT를 이용해 리뷰의 키워드 추출  Model : SBERT  형태소 분석 : OKT(OPEN KOREAN TEXT) 사용  키워드 추출 : Sklearn-CountVectorizer 사용  키워드 다양화 : max\_sum\_similarity | | 4 | 데이터 시각화 | 담당 : 정 희  시각화 모듈 : matplotlib, wordcloud   1. Bar\_graph      1. Wordcloud | | 5 | Ppt 및 포트폴리오 제작 | 담당 : 팀 전체 | | | | | | | | | |
| **상세개발내용** | |  | | | | | | |
|  | |  | | | | | | |
|  | | | | | | | | |
| **과제수행방법** | |  | | | | | | |
|  | | 1. Pyqt5 designer를 이용한 ui제작 2. 유저와 상호작용하는 시그널 함수 입력 3. 기능을 수행 하는 test코드 작성 4. test코드 모듈화 및 main 파일에 코드 작성 5. main파일 기능 수행 확인 | | | | | | |
| |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 주요내용 | 추 진 일 정 (일) | | | | | | | | | | | | | | | 12~16 | 19 | 20 | 21 | 22 | 23 | 26 | 27 | 28 | 29 | 30 | 1 | 2 | 3 | | 계획서 제작 및 설계 | ⊙ |  |  |  |  |  |  |  |  |  |  |  |  |  | | 데이터 수집 및 데이터베이스 설계 |  | ⊙ | ⊙ | ⊙ | ⊙ | ⊙ |  |  |  |  |  |  |  |  | | 상품평 분석 |  |  |  |  |  | ⊙ | ⊙ | ⊙ |  |  |  |  |  |  | | Front\_end연동 |  |  |  |  |  |  |  | ⊙ | ⊙ | ⊙ | ⊙ | ⊙ | ⊙ |  | | 시각화 |  |  |  |  |  |  |  |  |  | ⊙ | ⊙ |  |  |  | | 프로젝트 관련 문서제작 |  |  |  |  |  |  |  |  |  |  |  |  |  | ⊙ | | | | | | | | | |
| **최종결과물** | | |  | | | | | |
|  | | | 1. input page  1. main page | | | | | |
|  | | | | | | | | |
| **결과활용계획** | | |  | | | | | |
|  | | | 1. 리뷰에서 분석해낸 키워드를 통해 상품소개에서 제공하는 판매자의 주관적인 의견이 아니라 구매자의 입장에서의 상품의 평가를 확인 2. 상품의 실물을 보지못하고 구매하는 인터넷 쇼핑에서 먼저 상품을 구매한 구매자들의 상품평을 분석, 지금 구매하려는 구매자에게 정보를 제공 3. 인터넷 쇼핑에 실패할 확률을 줄임 | | | | | |
|  | | | | | | | | |
| **애로사항** | | |  | | | | | |
|  | | | 1. 계획서 제작 단계    1. 쿠팡 크롤링 차단 문제       1. 대안   쿠팡이 아닌 다른 사이트로 변경  우회 방법 찾기   * + 1. 해결 방안   쿠팡 홈페이지에서 자동화 된 소프트웨어 라는 정보를 획득하기 직전까지만 데이터 수집   * + 1. 문제점   양질의 데이터 수집 불가   1. 데이터 수집 및 데이터 베이스 설계   2-1 detail\_table의 효용성 문제  2-1-1) 대안  Prod\_table과 합치자  계획서 대로 진행  2-1-2) 해결방안  Prod\_\_table과 합쳐서 프로젝트 진행  2-2 review\_table 키워드 추출 문제  2-2-1) 대안  리뷰마다 키워드를 추출  모든 리뷰를 하나로 만들어 키워드 추출  2-2-2) 해결방안  리뷰마다 키워드를 추출해 프로젝트 진행  2-2-3) 문제점  키워드 추출 모듈 실행 시간이 매우 길어 짐  2-3 리뷰 키워드 추출 방식  2-3-1)의견  Input page에서 리뷰 수집 시 키워드 추출도 같이 진행  모든 리뷰를 데이터 베이스에 저장 후 main page에서 키워드 추출 진행  2-3-2) 해결방안  모든 리뷰를 데이터 베이스에 저장 후 main page에서 키워드 추출 하는 방식으로 프로젝트 진행   1. Front\_end   3-1 pyscript selenium모듈 사용 불가 문제  3-1-1)해결 방안  pyqt5로 Front\_end 진행   1. 시각화   4-1 계획했던 rader\_chart 기술부족으로 구현 불가 문제  4-1-1 ) 해결 방안  Wordcloud로 변경해 유저에게 정보 제공  4-2 계획했던 pie\_\_chart 기간부족으로 구현 실패 문제  4-2-1) 해결방안  구현 계획 취소 | | | | | |
|  | | | | | | | | |
|  | | | | | | | | |