

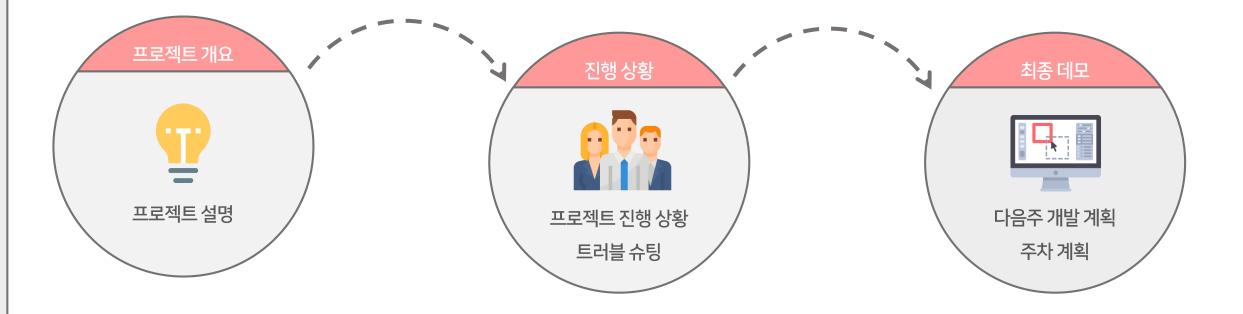
# 커뮤니티 상 거래를 위한 게시글 제목 변경 웹 확장 프로그램

팀명 : 제몫 박시현 + 박진영 + 이민희













#### 프로젝트 설명

♥무료배송♥ 워싱 면 라이더 자켓 판매 ❷ ◎

니트,필립림 자켓,양가죽 패딩코트(BCBG,Part or lady,마이클 코어) 판매 @ ◎

\*\* ♠♠ 셀린느 지고트 레니본 시스템 CK 존스뉴욕 아디다스 모르간 MNG ♠♠4

판매 🝛 [2] 🕲

워싱 면 라이더 자켓 | 15000 | 사용감 유 | 2주

필립림 자켓 | 30000 | 새상품 | 0회

존스뉴옥 | 43000 | 최상 | 착용만

#### 프로젝트 진행 상황

박시현

장고를 통한 웹서버 구축 및 연동

aws서버 구축 및 연동 / 데이터셋 추가 확보

박진영

프론트엔드 상의 통신 구현

프론트엔드 상 크롤링 구현, 프론트엔드 수정

이민희

모델 파라미터 변경

성능 및 정확도 향상

#### 데이터 셋 추가 확보

• • • • • • • • • • • • • •

- KT.txt
- LGU.txt
- SKT.txt
- ☐ 남성상의.txt
- ☐ 남성상의2.txt
- 남성상의3.txt
- 남성상의4.txt
- 남성신발.txt
- 남성신발2.txt
- ☐ 남성신발3.txt
- 남성하의.txt
- 남성하의2.txt
- 남성하의3.txt

- 노트북.txt
- 노트북2.txt
- 여성상의.txt
- 여성상의2.txt
- 여성상의3.txt
- 여성상의4.txt
- ☐ 여성신발.txt
- 여성신발2.txt
- 여성하의.txt
- 여성하의2.txt
- 태블릿PC.txt
- 태블릿PC2.txt
- 태블릿PC3.txt

여성상의 3만

남성상의 2만

남성신발 3만

노트북 2만

여성신발 1만

남성하의 1만

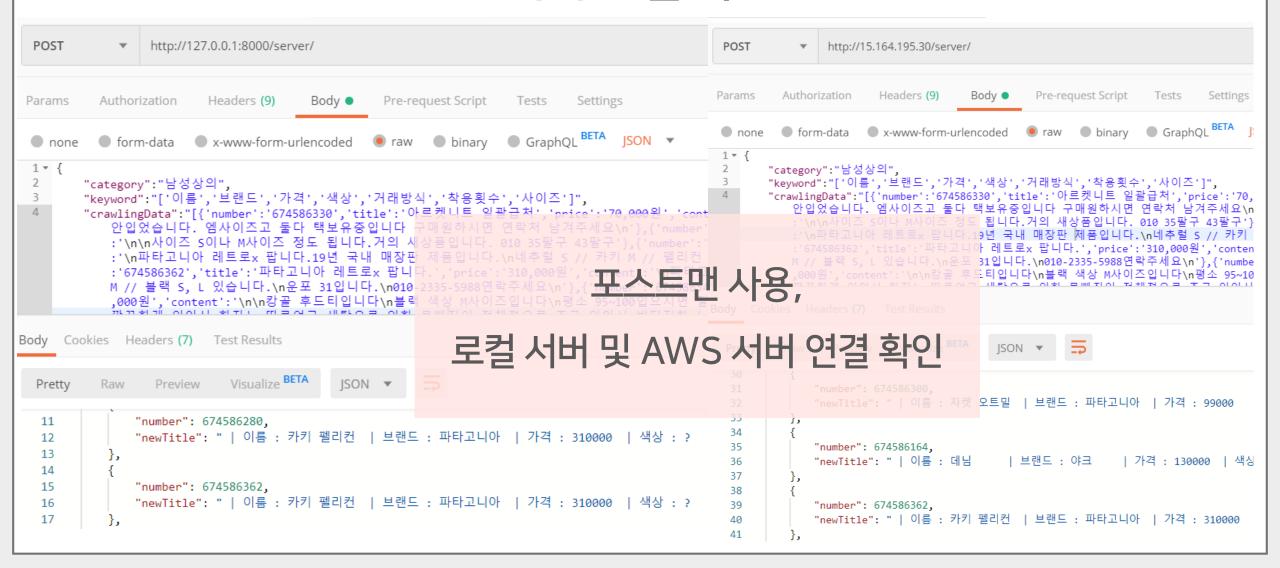
태블릿PC 2만

데이터 총 14만개 추가 확보



#### 서버 연결 확인

. . . . . . . . . . . . . . . .





#### 서버 연결 확인

. . . . . . . . . . . . . . . .

느린 크롤링 속도

리눅스 서버에서의 네이버 로그인 캡챠

=> 자바스크립트 크롤링을 통해 해결

(속도문제 해결, 로그인 필요 x)

#### 프론트엔드 상 크롤링

#### 콘텐트 스크립트 상에서 크롤링

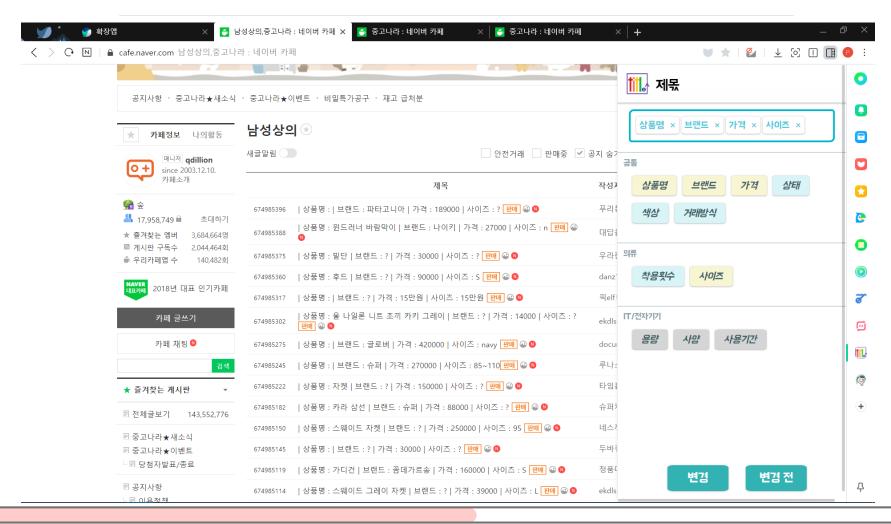
\$.get(url, function(data) 사용

동일 도메인 요청 시 세션쿠키 유지로 사용자 로그인 정보 사용

#### 비동기 처리로 속도 빠름

#### 프론트엔드 상 통신 연결

• • • • • • • • • • • • • • •











#### 프론트엔드 수정

. . . . . . . . . . . . . . . .



#### IT 용량 추출 방법

HDD/SSD 유무 파악



Min과 비슷한 숫자 중 최대값 반환



HDD/SSD 중 있는 단어와 매칭

#### IT 브랜드명

. . . . . . . . . . . . . . . .

- 1. 브랜드 명이 영어
- 2. 상품명 (영어, 숫자 포함)
- 3. 상세 스펙 나열

모든 경우 처리



#### 사전 생성

## 모든 사전 생성 완료 및 연결

수정한 날짜	유형	크기
2019-12-01 오후 8	텍스트 문서	6KB
2019-12-01 오후 8	텍스트 문서	5KB
2019-12-01 오후 6	텍스트 문서	1KB
2019-12-01 오전 1	텍스트 문서	1KB
2019-12-01 오후 6	텍스트 문서	1KB
2019-12-01 오후 8	텍스트 문서	7KB
2019-12-01 오후 8	텍스트 문서	5KB
2019-12-01 오전 1	텍스트 문서	6KB
2019-12-01 오전 1	텍스트 문서	2KB
	2019-12-01 오후 8 2019-12-01 오후 8 2019-12-01 오후 6 2019-12-01 오전 1 2019-12-01 오후 6 2019-12-01 오후 8 2019-12-01 오후 8 2019-12-01 오전 1	수정한 날짜

#### 원치 않는 값 후처리

• • • • • • • • • • • • • •

```
def getState(string list, docs list):
#상태 반환
find keyword = re.findall("\w+글", string list)
find keyword = list(filter(('취급').__ne__, find_keyword))
find keyword = list(filter(('발급'). ne , find keyword))
find keyword = list(filter(('자급'). ne , find keyword))
find keyword = list(filter(('동급'). ne , find keyword))
find keyword = list(filter((|X| = |X|)). ne , find keyword))
 if sel color == '':
  find keyword = re.findall("\w+ ?색",string_list)
  find keyword = list(filter(('배색').__ne__, find_keyword))
  find keyword = list(filter(('검색'). ne , find keyword))
  find_keyword = list(filter(('변색').__ne__, find_keyword))
```





# 최종데모

#### 다음주 개발 계획

박시현

박진영

이민희

최종레포트작성

#### 팀원 별 각 주간 개발 계획

••••••

박시현	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
주제 확정/제안서 작성															
크롤링 · 전처리															
데이터셋 구축								중간 고사							최종
서버 구축								& 중간						최 <u>종</u> 데모	최종 리포 트
통신 모듈 작업								데모 준비							제출
서버 통신 구현															
데모 준비ㆍ서비스 테스트															

#### 팀원 별 각 주간 개발 계획

••••••

박진영	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
주제 확정/제안서작성															
프론트엔드 화면 구현															
프론트엔드 동작 구현								중간 고사							최종
텍스트 대치 구현								& 중간						최종 데모	리포     트
확장 앱 통신 구현								데모 준비							제출
웹 페이지 쿠키 구현															
테스트															

## 팀원 별 각 주간 개발 계획

••••••

이민희	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
주제 확정/제안서작성															
자연어 처리 공부															
라이브러리 사용 전처리															
알고리즘 구체화								중간							
문장분리								고사 &						최종	최종 리포
주변 토큰 처리								중간 데모						데모	트 제출
키워드 사전 변경								.;. 준비							""
모델 구축, 파라미터 변경															
서버 모듈 연결															
성능 및 정확도 향상															

#### 전체 각 주간 개발 계획

••••••

공통: 연두색 / 박시현:핑크색 / 박진영:하늘색 / 이민희:주황색

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
주제확정/제안서 작성																	
크롤링/데이터 확보																	
데이터 전처리																	
프론트 엔드								중간							417		
서버 구축								고사 &						최종	최종		최종 리포
텍스트 대치 구현								중간 데모						데모	트 제출		
자연어 처리 모델 구축								준비							"=		
웹 페이지 쿠키 구현																	
모듈 연결																	
유닛 · 전체 테스트																	



