

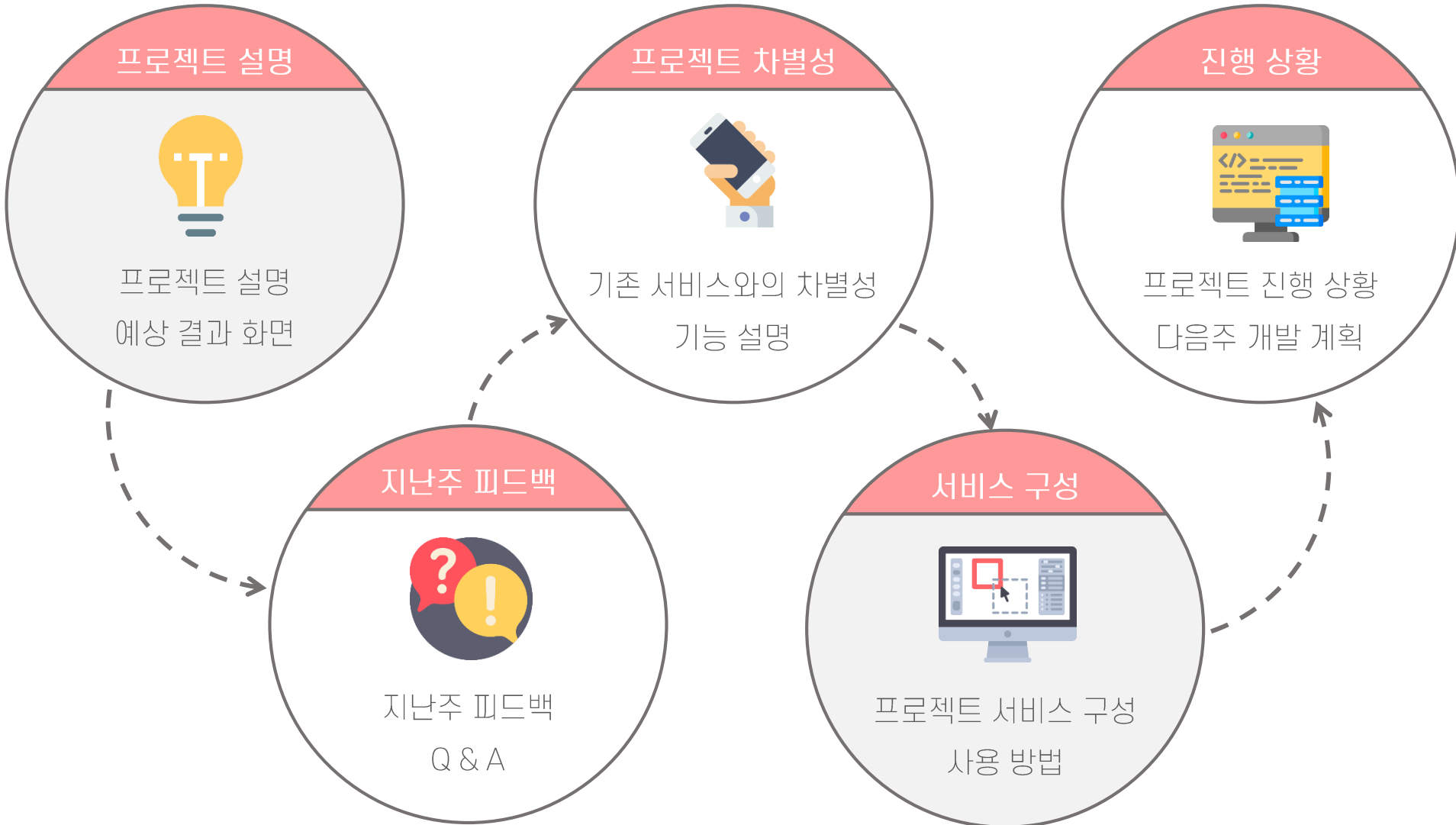


커뮤니티 상 거래를 위한 게시글 제목 변경 웹 확장 프로그램



팀명 : 제뭇
박시현 + 박진영 + 이민희

TABLE OF CONTENTS



프로젝트 설명

♥ 무료배송 ♥ 워싱 면 라이더 자켓 판매 🧑🏻 N

니트, 필립 림 자켓, 양가죽 패딩코트(BCBG, Part or lady, 마이클 코어) 판매 🧑🏻 N

** ♠️ ♠️ 셀린느 지고트 레니본 시스템 CK 존스뉴욕 아디다스 모르간 MNG ♠️ ♠️ 4
판매 🧑🏻 [2] N



워싱 면 라이더 자켓 | 15000 | 사용감 유 | 2주

필립 림 자켓 | 30000 | 새상품 | 0회

존스뉴욕 | 43000 | 최상 | 착용만

지난주 Feedback 및 Q&A

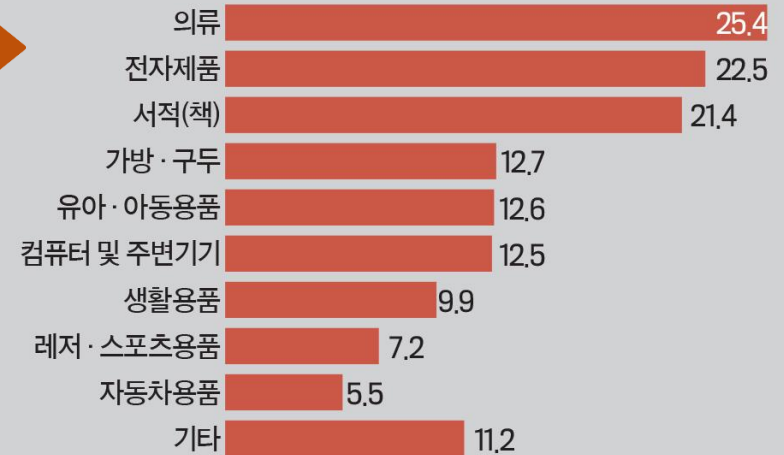
Q. 어느 정도 커버를 할 수 있을지?

A. IT기기와 옷에 대해 카테고리 분류



최근 1년 동안 중고거래 많은 제품

단위: %



* 중고거래 이용자 657명 기준(중복응답), 자료: 한국리서치

Q. 게시물 하나하나 들어갈 때 걸리는 속도 고려

A. Selenium 사용시 시간이 걸리는 것을 확인, Bea

Q. 네이버 어뷰징 제제에 걸릴 가능성?

A. 네이버 로그인 정보를 가져와서 사용할 예정

중고나라 연간 방문자 수

단위: 명



지난주 Feedback 및 Q&A

Q. 어느 정도 커버를 할 수 있을지?

A. IT기기와 옷에 대해 카테고리 분류

Q. 게시물 하나하나 들어갈 때 걸리는 속도 고려?

A. Selenium 사용시 시간이 걸리는 것을 확인, BeautifulSoup으로 크롤링 진행

Q. 네이버 어뷰징 제재에 걸릴 가능성?

A. user-urgent 값 추가 또는 time.sleep 사용

차별성

보나파이드

- 카페에서 지정한 형식으로만 제목 설정 가능
- 모든 카페에 대해서 적용되는 것이 아님

새로운 서비스

- 사용자가 직접 항목을 선택해 선택된 항목에 대해서만 게시글 제목이 변환됨
- 판매자가 게시물 제목 작성 시 불편함을 피할 수 있음
- 네이버나 다음 카페에서 모두 적용 가능

프로젝트 구성

★ 즐겨찾는 멤버

게시판 구독수

우리카페앱 수

3,682,524명

2,067,236회

139,663회

NAVER 대표카페

2018년 대표 인기카페

카페 글쓰기

카페 채팅

검색

★ 즐겨찾는 게시판

전체글보기

중고나라 ★새소식

중고나라 ★이벤트

당첨자발표/종료

공지사항

이용정책

중고나라 이야기

물거래 후기(칭찬)

어려운 이웃 사연

중고나라 행복나눔

제품 무료나눔

먹을거리 무료나눔

손글씨/이미지나눔

제목

작성자

작성일

조회

643645581

빅사이즈 밍크코트 밍크카라코트 여러벌 정리중 [판매]

조금씩비움

01:08

1

643645258

막스마라캐시미어술, 타원라마코트, 발렌시아캐시미어코트 // 타임폭스퍼패딩 자켓 앤블레트 바네사부르노코트, 구오, 마쥬 피어리패딩 에스다밍크 ~~겨울의류 [판매]

han스타일

01:06

6

643645090

가을 아우터 만원부터~ 마인 율가디건/쉬즈마스 트위드 자켓/트렌치 코트/로스 야상/간절기 후드집업/모르간/쿠아/린넨 코트 데님 더블 포켓 jk/디자인 유니이티드 항공점퍼/아디다스 슈퍼스타 트랙탑 저지 [판매]

탄산수덕후

01:05

5

643645062

코발트 여성 가죽자켓 사이즈77 15000원에 팝니다 [판매]

화사한혜진

01:05

2

643644655

** ♣♣ 미사원피스 ASOS원피스 SJSJ스커트 외 ♣♣5 [판매]

타미타모

01:02

3

643644202

♥무료배송♥ 워싱 면 라이더 자켓 [판매]

hkk14

01:00

2

643643803

니트,필립 린 자켓,양가죽 패딩코트(BCBG,Part or lady,마이클 코어) [판매]

powerbeat97

00:58

1

643643625

** ♣♣ 셀린느 지고트 레니본 시스템 CK 존스뉴욕 아디다스 모르간 MNG ♣♣4 [판매]

타미타모

00:57

1

643643611

오즈세컨 자켓,후드점퍼 [판매]

그럼에도봄

00:57

5

643643474

블랙업 보이프렌드 자켓 새상품 운포 2만원 [판매]

youthxyouth

00:56

5

643643272

타임 . 마인 . 마쥬 . 시스템 캐시미어 [판매]

MORDEN109

00:55

20

643643210

레니본 코트 [판매]

per

00:55

6

643643032

올리브데올리브 예쁜점퍼자켓 [판매]

그럼에도봄

00:54

2

643642988

지센 양가죽자켓 XL사이즈 2만원에팝니다 [판매]

화사한혜진

00:54

4

643642917

플로 탈프로렌 자켓 [판매]

ppong5624

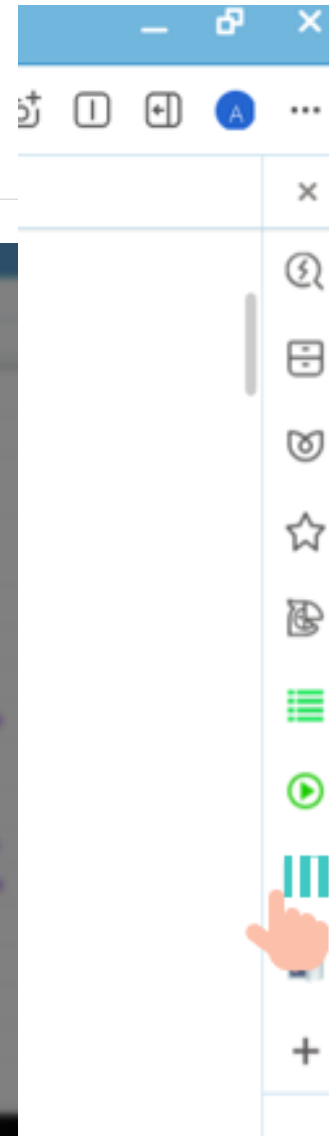
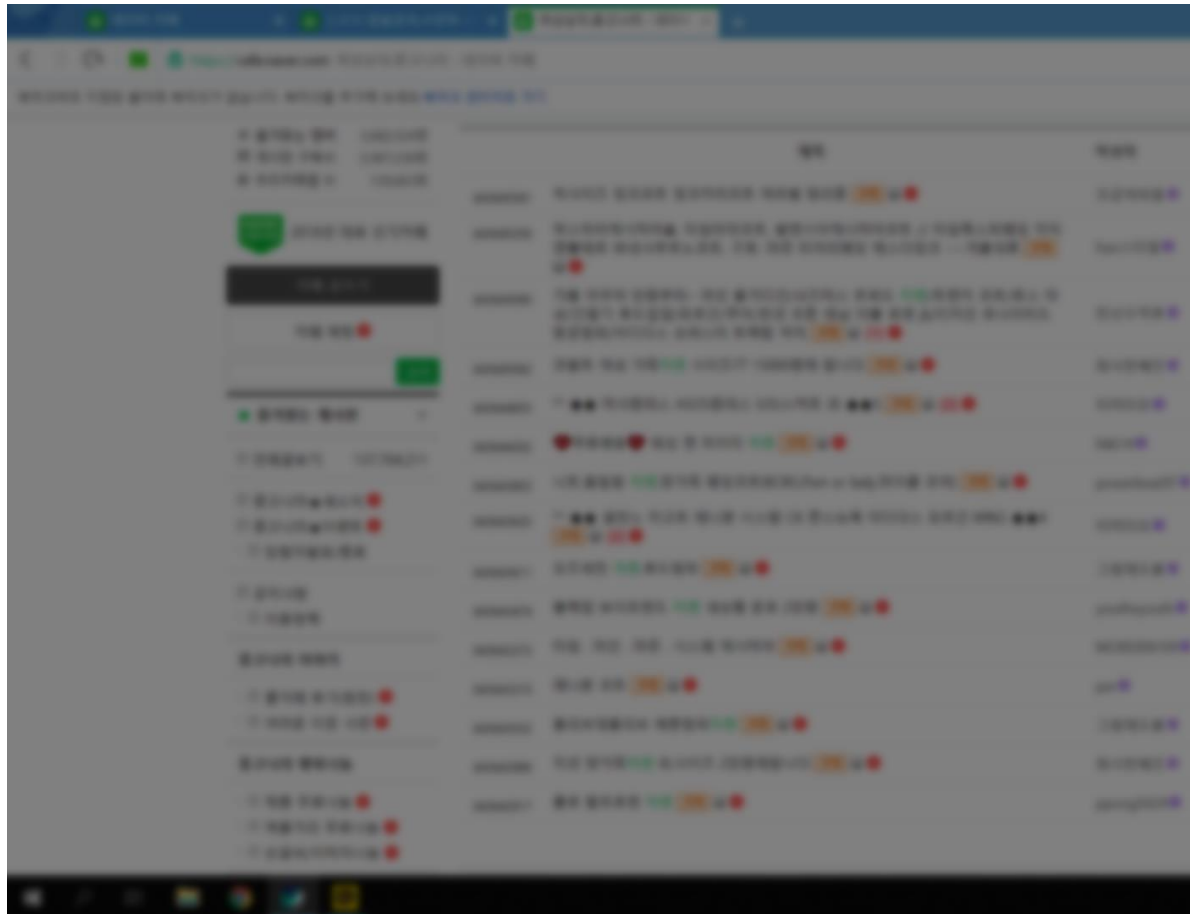
00:54

3

오전 1:26

2019-09-24

프로젝트 구성



확장 앱 아이콘 클릭

프로젝트 구성

The screenshot shows a web browser window with the address bar displaying 'https://cafe.naver.com 여성상의,중고나라 : 네이버 카페'. The page content includes a sidebar on the left with navigation links like '★ 즐겨찾는 멤버', '2018년 대표 인기카페', and '카페 글쓰기'. The main area displays a list of products with columns for '제목' (Title) and '작성자' (Author). A filter sidebar on the right is open, showing '선택된 제목 형태' (Selected Title Format) with '이름 x' selected, and '카테고리 (키워드 적용 상품)' (Category (Keyword Applied Product)) with '기본' (Basic) selected. The '기본' category has sub-filters for '이름' (Name), '가격' (Price), '브랜드' (Brand), '상태' (Status), '실사용기간' (Actual Usage Period), '거래방식' (Transaction Method), and '색상' (Color). The '이름' filter is currently selected. The bottom of the browser window shows the Windows taskbar with the date '2019-09-24' and time '오전 1:25'.

번호	제목	작성자
643645581	빅사이즈 밍크코트 밍크카라코트 여러벌 정리중 [판매] [1]	조금씩비움
643645258	막스마라캐시미어술, 타일라마코트, 발렌시아캐시미어코트 // 타일폭스퍼패딩 자켓 연봉테르 바네사부르노코트, 구로. 마쥬 띠어리패딩 에스다밍크 ~~겨울의류 [판매] [1]	han스타일
643645090	가을 아우터 만원부터~ 마인 올라다건/쉬즈미스 트윈드 자켓/트렌치 코트/로스 야상/간절기 후드집업/모르간/쿠아/린넨 코트 데님 더블 포켓 jk/디자인 유니타이트 항공점퍼/아디다스 슈퍼스타 트랙탑 저지 [판매] [1]	탄산수덕후
643645062	코발트 여성 가죽자켓 사이즈77 15000원에 팝니다 [판매] [1]	화사한혜진
643644655	** ♣♣ 미사원피스 ASOS원피스 SJSJ스커트 외 ♣♣5 [판매] [2]	타미타모
643644202	♥무료배송♥ 워싱 면 라이더 자켓 [판매] [1]	hkk14
643643803	니트,필립 림 자켓,왕가죽 패딩코트(BCBG,Part or lady,마이클 코어) [판매] [1]	powerbeat97
643643625	** ♣♣ 셀린느 지고트 레니본 시스템 CK 존스뉴욕 아디다스 모르간 MNG ♣♣4 [판매] [2]	타미타모
643643611	오즈세션 자켓,후드점퍼 [판매] [1]	그럼에도봄
643643474	블랙업 보이프렌드 자켓 생상품 운포 2만원 [판매] [1]	youthxyouth
643643272	타일 . 마인 . 마쥬 . 시스템 캐시미어 [판매] [1]	MORDEN105
643643210	레니본 코트 [판매] [1]	per
643643032	올리브 데올리브 예쁜점퍼자켓 [판매] [1]	그럼에도봄
643642988	지센 왕가죽자켓 XL사이즈 2만원에팝니다 [판매] [1]	화사한혜진
643642917	폴로 말프로렌 자켓 [판매] [1]	ppong5624

프로젝트 구성

우측 사이드바에서 키워드 선택

선택된 제목 형태

이름 x

카테고리 (키워드 적용 상품)

기본

이름

가격

브랜드

상태

사용기간

거래방식

색상

IT 기기 / 컴퓨터

약정

용량

부품 유무

음식 / 화장품

용량

유통기한

변경

프로젝트 구성

The screenshot shows a Naver Cafe page titled '여성상의,중고나라 : 네이버 카페'. The page displays a list of items for sale, including clothing and accessories. The sidebar on the left contains various filters and a search bar. The main content area shows a list of items with their titles, prices, and authors. A right-hand panel shows filters for '선택된 제목 형태' (Selected Title Format) and '카테고리 (키워드 적용 상품)' (Category (Keyword Applied Product)).

선택된 제목 형태

이름 x | 가격 x | 상태 x | 사용

카테고리 (키워드 적용 상품)

기본

이름 | 가격 | 브랜드

상태 | 실사용기간

거래방식 | 색상

IT 기기 / 컴퓨터

약정 | 용량 | 부품 유무

음식 / 화장품

용량 | 유통기한

변경

제목

643645581 빅사이즈 밍크코트 밍크카라코트 여러벌 정리중 [판매] [1]

643645258 막스마라캐시미어술, 타임라마코트, 발렌시아캐시미어코트 // 타임폭스퍼패딩 자켓 연봉테르 바네사부르노코트, 구로, 마쥬 피어리패딩 에스다핑크 ~~겨울의류 [판매]

643645090 가을 아우터 만원부터~ 마인 올가디건/쉬즈미스 트윈드 자켓/트렌치 코트/로스 야상/간절기 후드집업/모르간/쿠아/린넨 코튼 데님 더블 포켓 jk/디자인 유니타이트 항공정퍼/아디다스 슈퍼스타 트랙탑 저지 [판매] [1]

643645062 코발트 여성 가죽자켓 사이즈77 15000원에 팝니다 [판매]

643644655 ** ♣♣ 미사원피스 ASOS원피스 SJSJ스커트 외 ♣♣5 [판매] [2]

643644202 ♥무료배송♥ 워싱 면 라이더 자켓 [판매]

643643803 니트,필립 린 자켓,왕가죽 패딩코트(BCBG,Part or lady,마이클 코어) [판매]

643643625 ** ♣♣ 셀린느 지고트 레니본 시스템 CK 존스뉴욕 아디다스 모르간 MNG ♣♣4 [판매] [2]

643643611 오즈세션 자켓,후드점퍼 [판매]

643643474 블랙업 보이프렌드 자켓 생상품 운포 2만원 [판매]

643643272 타임, 마인, 마쥬, 시스템 캐시미어 [판매]

643643210 레니본 코트 [판매]

643643032 올리브데올리브 예쁜점퍼자켓 [판매]

643642988 지센 왕가죽자켓 XL사이즈 2만원에팝니다 [판매]

643642917 풀로 말프로렌 자켓 [판매]

작성자가

조금씩비움

hans타일

탄산수덕후

화사한혜진

타미타모

hkk14

powerbeat97

타미타모

그림에도봄

youthxyouth

MORDEN109

per

그림에도봄

화사한혜진

ppong5624

★ 즐겨찾는 멤버 3,682,524명

게시판 구독수 2,067,236회

우리카페업 수 139,663회

2018년 대표 인기카페

카페 글쓰기

카페 채팅

검색

★ 즐겨찾는 게시판

전체글보기 137,704,211

중고나라★새소식

중고나라★이벤트

당첨자발표/종료

공지사항

이용정책

중고나라 이야기

즐거워 후기(칭찬)

어려운 이웃 사연

중고나라 행복나눔

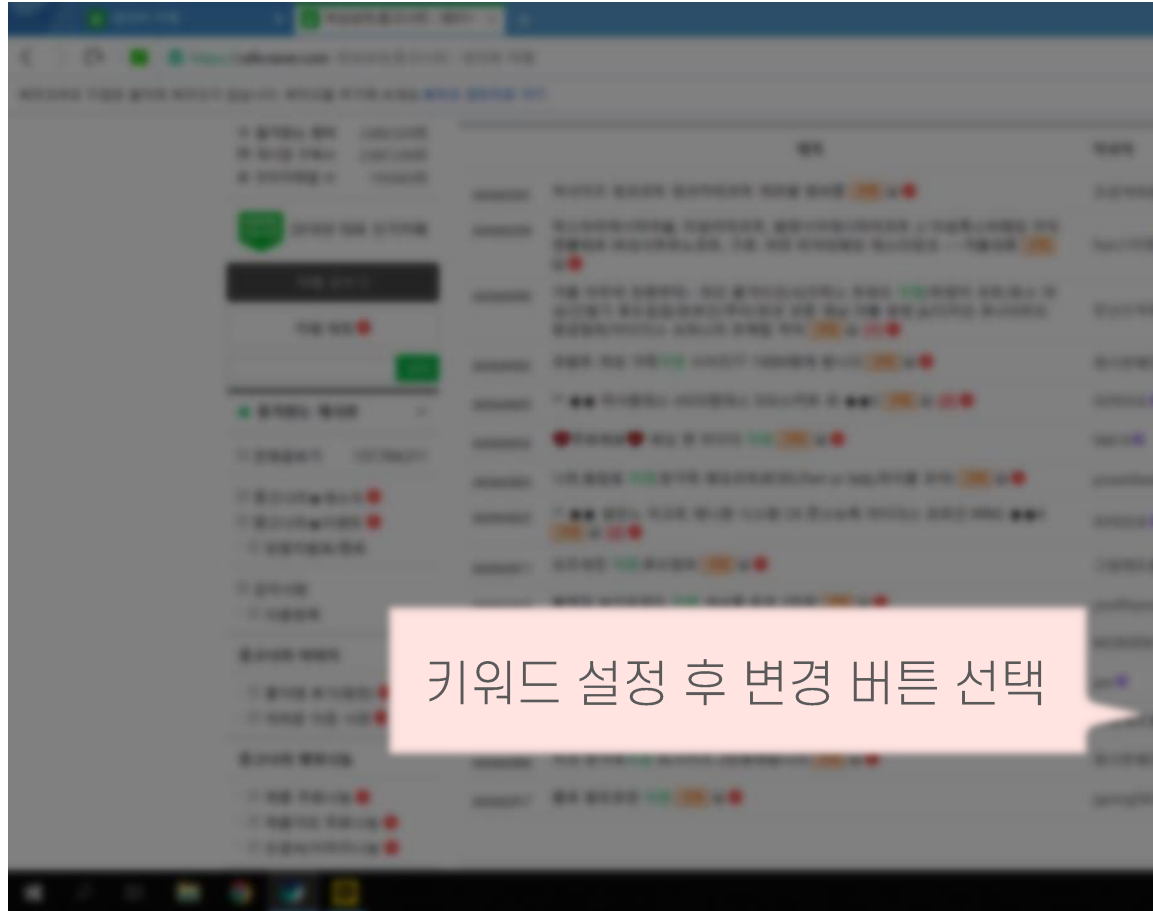
제품 무료나눔

먹을거리 무료나눔

손글씨/이미지나눔

오전 1:25
2019-09-24

프로젝트 구성



키워드 설정 후 변경 버튼 선택

선택된 제목 형태

이름 x | 가격 x | 상태 x | 사용

카테고리 (키워드 적용 상품)

기본

이름 | 가격 | 브랜드

상태 | 사용기간

거래방식 | 색상

IT 기기 / 컴퓨터

약정 | 용량 | 부품 유무

음식 / 화장품

용량 | 유통기한

변경

프로젝트 구성

네이버 카페 | 수시 합격결과,수만원-수 | 여성상의,중고나라 : 네이버 | +

https://cafe.naver.com 여성상의,중고나라 : 네이버 카페

북마크바로 지정된 폴더에 북마크가 없습니다. 북마크를 추가해 보세요. 북마크 관리자로 가기

★ 즐겨찾는 멤버 3,682,524명
게시판 구독수 2,067,236회
우리카페앱 수 139,663회

2018년 대표 인기카페

카페 글쓰기

카페 채팅

검색

★ 즐겨찾는 게시판

전체글보기 137,704,211

중고나라★새소식
중고나라★이벤트
당첨자발표/종료

공지사항
이용정책

중고나라 이야기

물거래 후기(칭찬)
어려운 이웃 사연

중고나라 행복나눔

재중 무료나눔
먹을거리 무료나눔
손글씨/이미지나눔

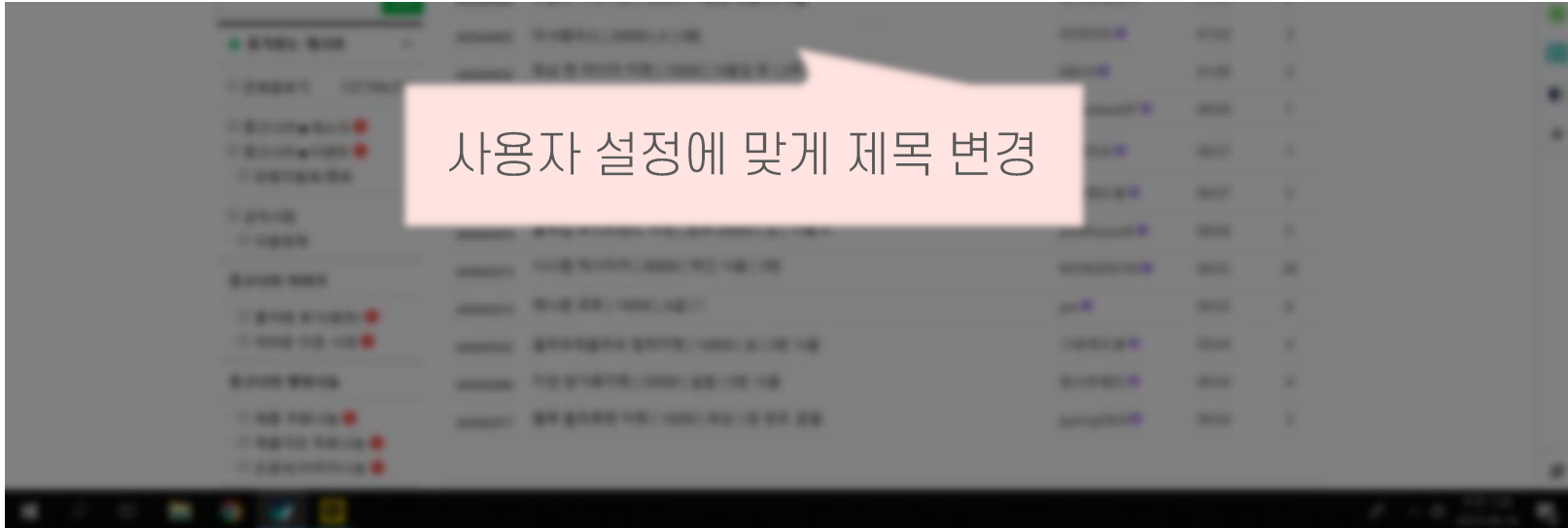
	제목	작성자	작성일	조회
643645581	링크 자켓 36000 양호 1년 전 구입	조금씩바움	01:08	1
643645258	캐시미어술 42000 최상 3개월	han스타일	01:06	6
643645090	쉬즈미스 트위드 자켓 25000 실착X 2회	탄산수덕후	01:05	5
643645062	코발트 가죽자켓 10000 사용감 있음 3개월	화사한혜진	01:05	2
643644655	미사원피스 20000 A 0회	타미타모	01:02	3
643644202	워싱 면 라이다 자켓 15000 사용감 유 2주	hkk14	01:00	2
643643803	폴립립 자켓 30000 새상품 0회	powerbeat97	00:58	1
643643625	존스뉴욕 43000 최상 착용만	타미타모	00:57	1
643643611	오즈세컨 자켓 12000 소매 오염 3회	그럼에도봄	00:57	5
643643474	블랙업 보이프렌드 자켓 운포 20000 상 사용 X	youthxyouth	00:56	5
643643272	시스템 캐시미어 38000 약간 사용 3번	MORDEN109	00:55	20
643643210	레니본 코트 15000 A급 ?	per	00:55	6
643643032	올리브데올리브 점퍼자켓 10000 상 3번 사용	그럼에도봄	00:54	2
643642988	자센 양가죽자켓 25000 실발 5번 사용	화사한혜진	00:54	4
643642917	플로 발프로렌 자켓 15000 최상 한 번도 없음	ppong5624	00:54	3

오전 1:26
2019-09-24

프로젝트 구성



643644202	워싱 면 라이더 자켓 15000 <u>사용감 유</u> 2주	hkk14	01:00	2
643643803	<u>필립 림</u> 자켓 30000 <u>새상품</u> 0회	powerbeat97	00:58	1
643643625	<u>존스뉴욕</u> 43000 최상 착용만	타미타모	00:57	1



프로젝트 구성

네이버 카페 | 수시 한불결과,수만원-수 | 여성상의,중고나라 : 네이버 | +

https://cafe.naver.com 여성상의,중고나라 : 네이버 카페

북마크바로 지정된 폴더에 북마크가 없습니다. 북마크를 추가해 보세요. 북마크 관리자로 가기

★ 즐겨찾는 열버 3,682,524명
게시판 구독수 2,067,236회
우리카페앱 수 139,663회

2018년 대표 인기카페

카페 글쓰기

카페 채팅

검색

★ 즐겨찾는 게시판

전체글보기 137,704,211

중고나라★새소식
중고나라★이벤트
당첨자발표/종료

공지사항
이용정책

중고나라 이야기

물거래 후기(칭찬)
어려운 이웃 사연

중고나라 행복나눔

재중 무료나눔
먹을거리 무료나눔
손글씨/이미지나눔

	제목	작성자	작성일	조회
643645581	링크 자켓 36000 양호 1년 전 구입	조금씩비용	01:08	1
643645258	캐시미어술 42000 최상 3개월	han스타일	01:06	6
643645090	쉬즈미스 트위드 자켓 25000 실착X 2회	탄산수덕후	01:05	5
643645062	코발트 가죽자켓 10000 사용감 있음 3개월	화사한혜진	01:05	2
643644655	미사원피스 2 코발트 여성 가죽자켓 사이즈 77 15000원에 팝니다		01:02	3
643644202	워싱 면 라이더 자켓 15000 사복집 뉴 2주	hkk14	01:00	2
643643803	폴립 자켓 30000 새상품 0회	powerbeat97	00:58	1
643643625	존스뉴욕 43000 최상 착용만	타미타모	00:57	1
643643611	오즈세컨 자켓 12000 소매 오염 3회	그럼에도봄	00:57	5
643643474	블랙업 보이프렌드 자켓 운포 20000 상 사용 X	youthx youth	00:56	5
643643272	시스템 캐시미어 38000 약간 사용 3번	MORDEN109	00:55	20
643643210	레니본 코트 15000 A급 ?	per	00:55	6
643643032	올리브데올리브 점퍼자켓 10000 상 3번 사용	그럼에도봄	00:54	2
643642988	자센 양가죽자켓 25000 실발 5번 사용	화사한혜진	00:54	4
643642917	플로 말프로렌 자켓 15000 최상 한 번도 없음	ppong5624	00:54	3

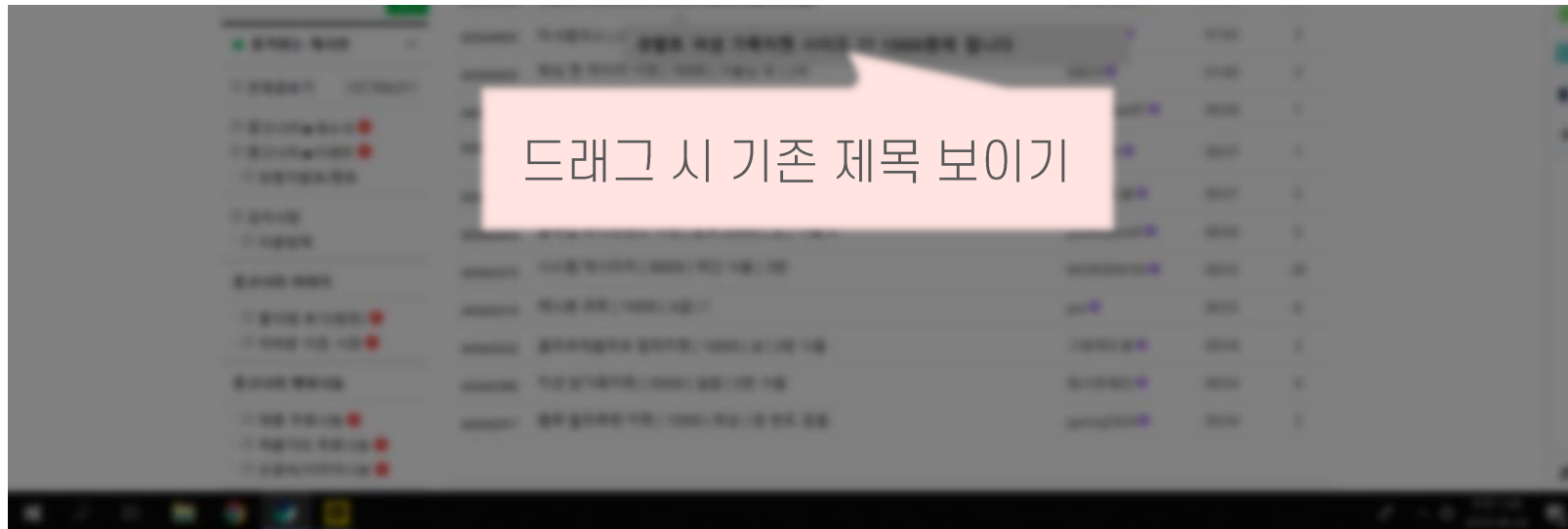
오전 1:26
2019-09-24

프로젝트 구성



643645062	코발트 가죽자켓 10000 사용감 있음 3개월	화사한혜진	01:05	2
643644655	미샤원피스 20000 사용감 있음 3개월	hkk14	01:02	3
643644202	워싱 면 라이더 자켓 15000 사용감 있음 2주	hkk14	01:00	2

코발트 여성 가죽자켓 사이즈 77 15000원에 팝니다



드래그 시 기존 제목 보이기

프로젝트 구성

The screenshot shows a Naver Cafe page titled '여성상의,중고나라 : 네이버 카페'. The page displays a list of items for sale, including clothing and accessories, with details like item ID, title, price, and seller. A sidebar on the right allows filtering items by category, price, and other attributes. The bottom of the page shows the Windows taskbar with various application icons and the system clock indicating 1:25 AM on September 24, 2019.

번호	제목	작성자
643645581	빅사이즈 밍크코트 밍크카라코트 여러벌 정리중 [판매]	조금씩비움
643645258	막스마라캐시미어술, 타원라마코트, 발렌시아캐시미어코트 // 타임폭스퍼패딩 자켓 엔볼트르 바네사부르노코트, 구호, 마쥬 띠어리패딩 에스다잉크 ~~겨울의류 [판매]	han스타일
643645090	가을 아우터 만원부터 ~ 마인 울가디건/위즈미즈 트윈드 자켓/트렌치 코트/로스 야 상/간절기 후드집업/모르간/쿠아/빈넨 코튼 데님 더블 포켓 jk/디자인 유나이티드 항공점퍼/아디다스 슈퍼스타 트랙탑 저지 [판매] [1]	탄산수덕후
643645062	코발트 여성 가죽자켓 사이즈77 15000원에 팝니다 [판매]	화사한혜진
643644655	** ♣♣ 미사원피스 ASOS원피스 SJS스커트 외 ♣♣5 [판매] [2]	타미타모
643644202	♥무료배송 ♥ 워싱 면 라이더 자켓 [판매]	hkk14
643643803	니트,필립 린 자켓,양가죽 패딩코트(BCBG,Part or lady,마이클 코어) [판매]	powerbeat97
643643625	** ♣♣ 셀린느 지고트 레니본 시스템 CK 존스뉴욕 아디다스 모르간 MNG ♣♣4 [판매] [2]	타미타모
643643611	오즈세컨 자켓,후드점퍼 [판매]	그럼에도봄
643643474	블랙업 보이프렌드 자켓 새상품 윤포 2만원 [판매]	youthxyouth
643643272	타임 . 마인 . 마쥬 . 시스템 캐시미어 [판매]	MORDEN105
643643210	레니본 코트 [판매]	per
643643032	올리브데올리브 예쁜점퍼자켓 [판매]	그럼에도봄
643642988	자센 양가죽자켓 XL사이즈 2만원에팝니다 [판매]	화사한혜진
643642917	폴로 말프로렌 자켓 [판매]	ppong5624

선택된 제목 형태

카테고리 (키워드 적용 상품)

기본

이름 가격 브랜드

상태 사용기간

거래방식 색상

IT 기기 / 컴퓨터

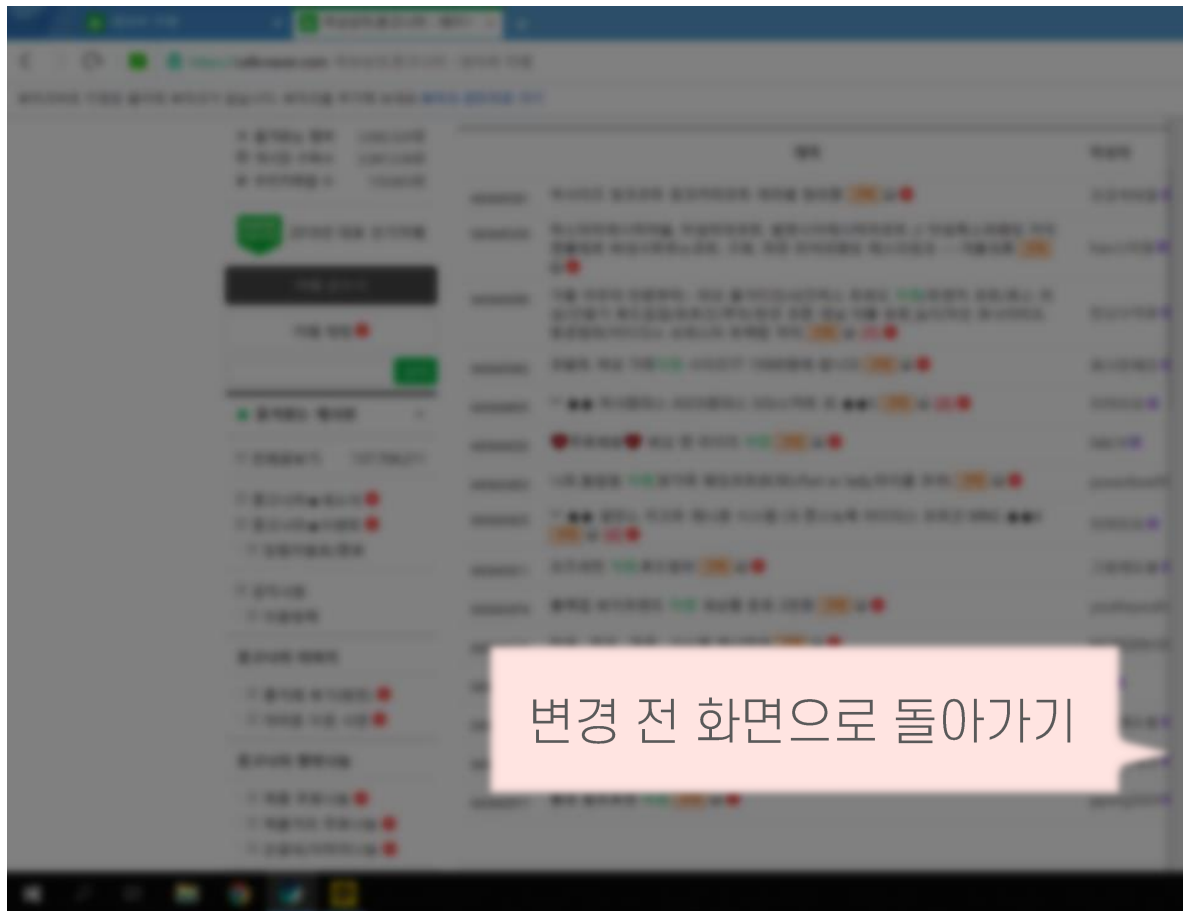
약정 용량 부품 유무

음식 / 화장품

용량 유통기한

변경 전

프로젝트 구성



변경 전 화면으로 돌아가기

선택된 제목 형태

카테고리 (키워드 적용 상품)

기본

이름

가격

브랜드

상태

사용기간

거래방식

색상

IT 기기 / 컴퓨터

약정

용량

부품 유무

음식 / 화장품

용량

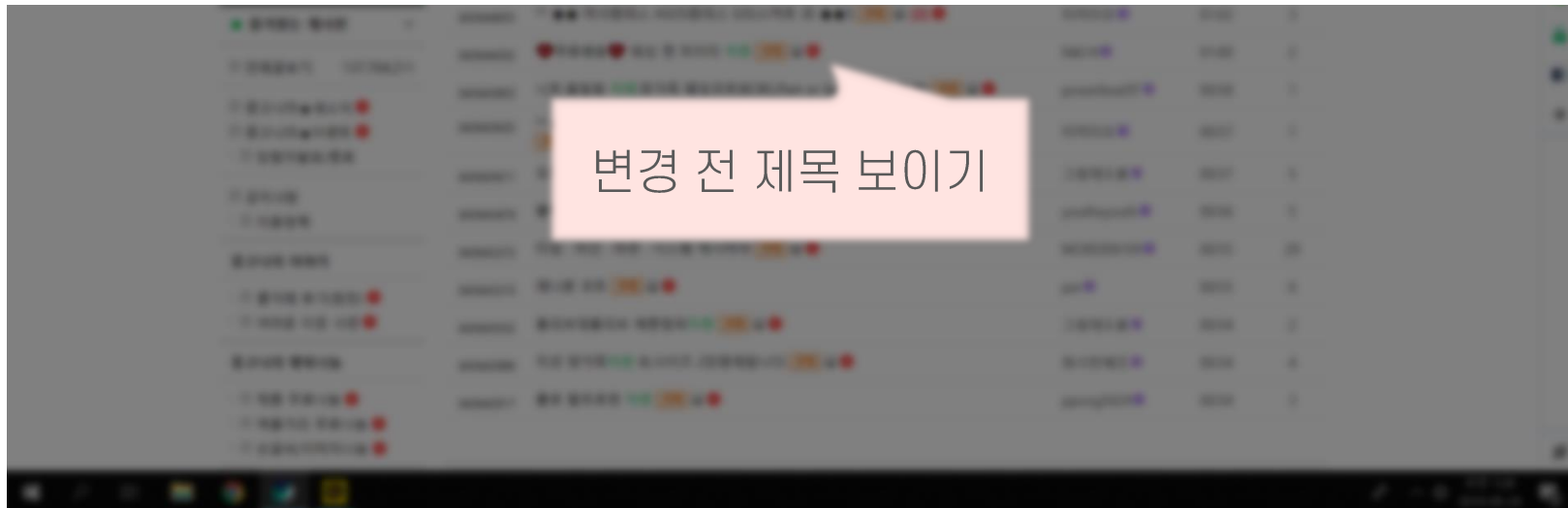
유통기한

변경 전

프로젝트 구성



643645062	코발트 여성 가죽자켓 사이즈77 15000원에 팝니다 판매 🌐 📌	화사한혜진 📌	01:05	2
643644655	** ♠️♠️ 미샤원피스 ASOS원피스 SJSJ스커트 외 ♠️♠️5 판매 🌐 [2] 📌	타미타모 📌	01:02	3
643644202	❤️무료배송❤️ 워싱 면 라이더 자켓 판매 🌐 📌	hkk14 📌	01:00	2



변경 전 제목 보이기

프로젝트 진행 상황

박시현

크롤링 연습 (Beautiful Soup 사용)
카페 html 구성 태그 분석

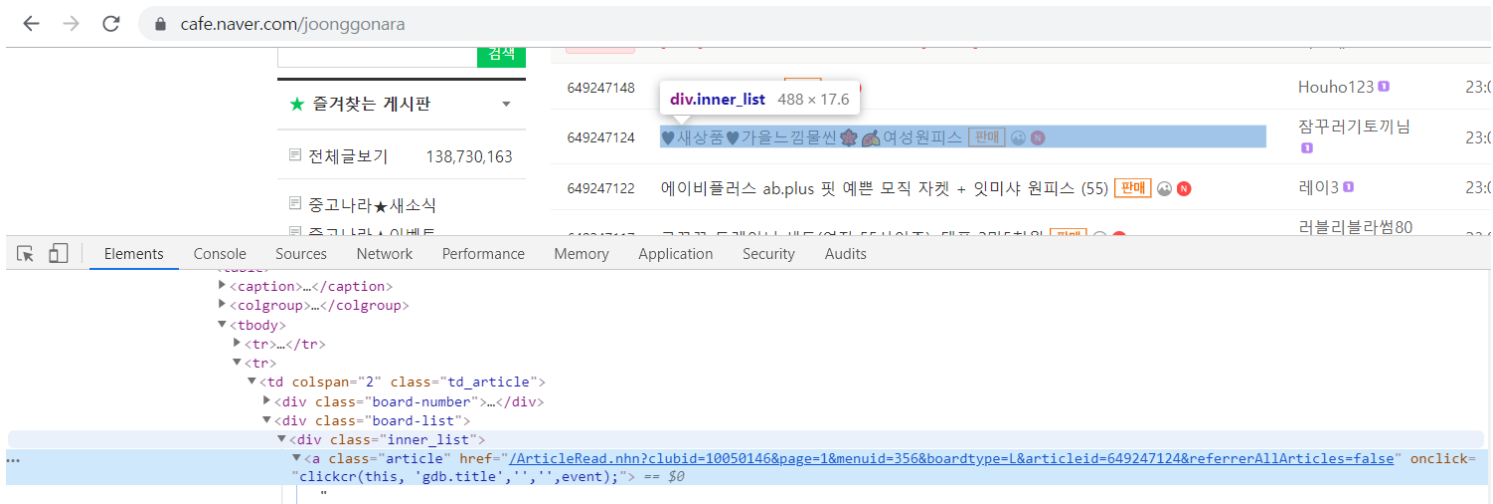
박진영

웹 프론트엔드
임시 icon 생성 및 side bar 띄우기

이민희

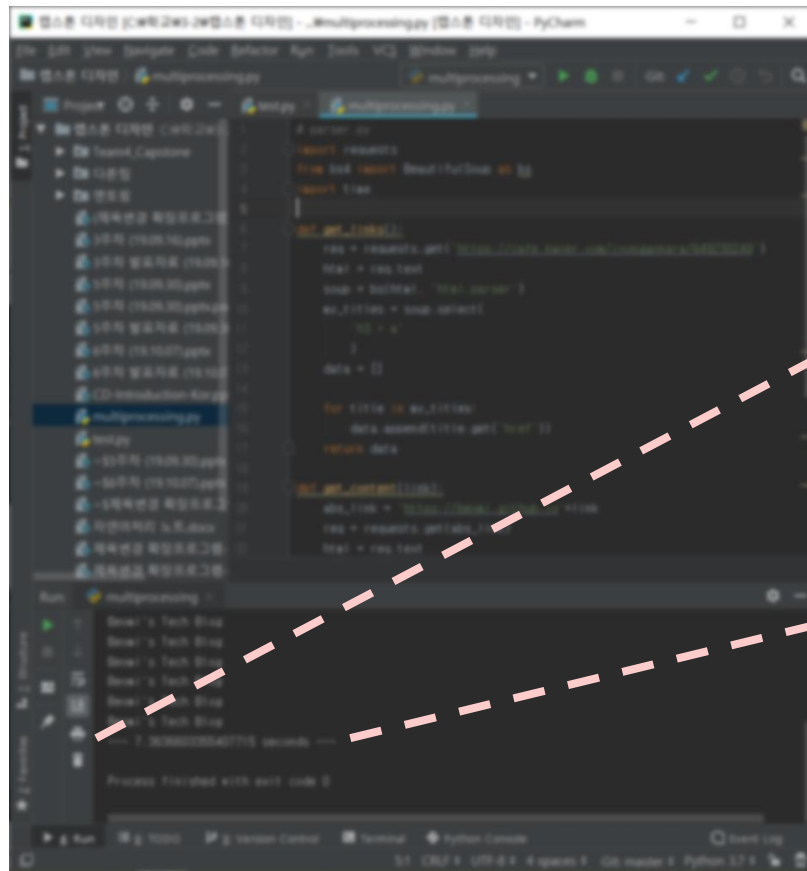
API 사용해서 자연어처리 연습 (KoNLPy, NLTK, Gensim)
데이터 토큰화 및 문서 벡터화 연습

html 태그 분석



페이지 소스 참고 후
주로 <table> 태그의 <a> 태그 사용 예정

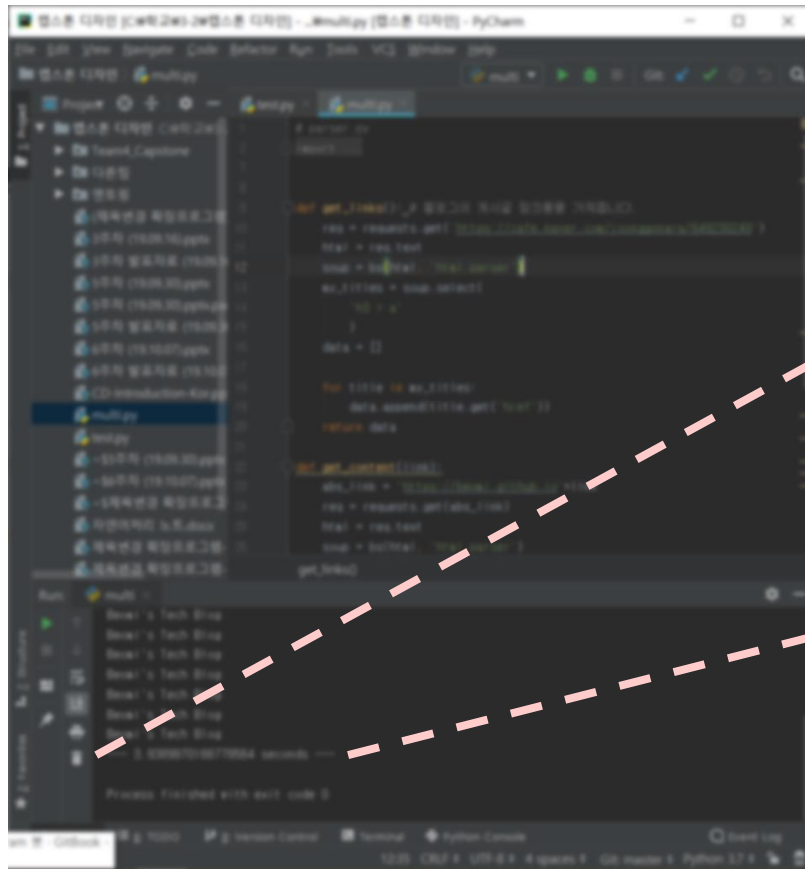
데이터 크롤링



=== 7.3636603355407715 seconds ===

약 25개 게시물 크롤링 시
대략 7초의 시간이 소요되는 것을 확인

Multiprocessing

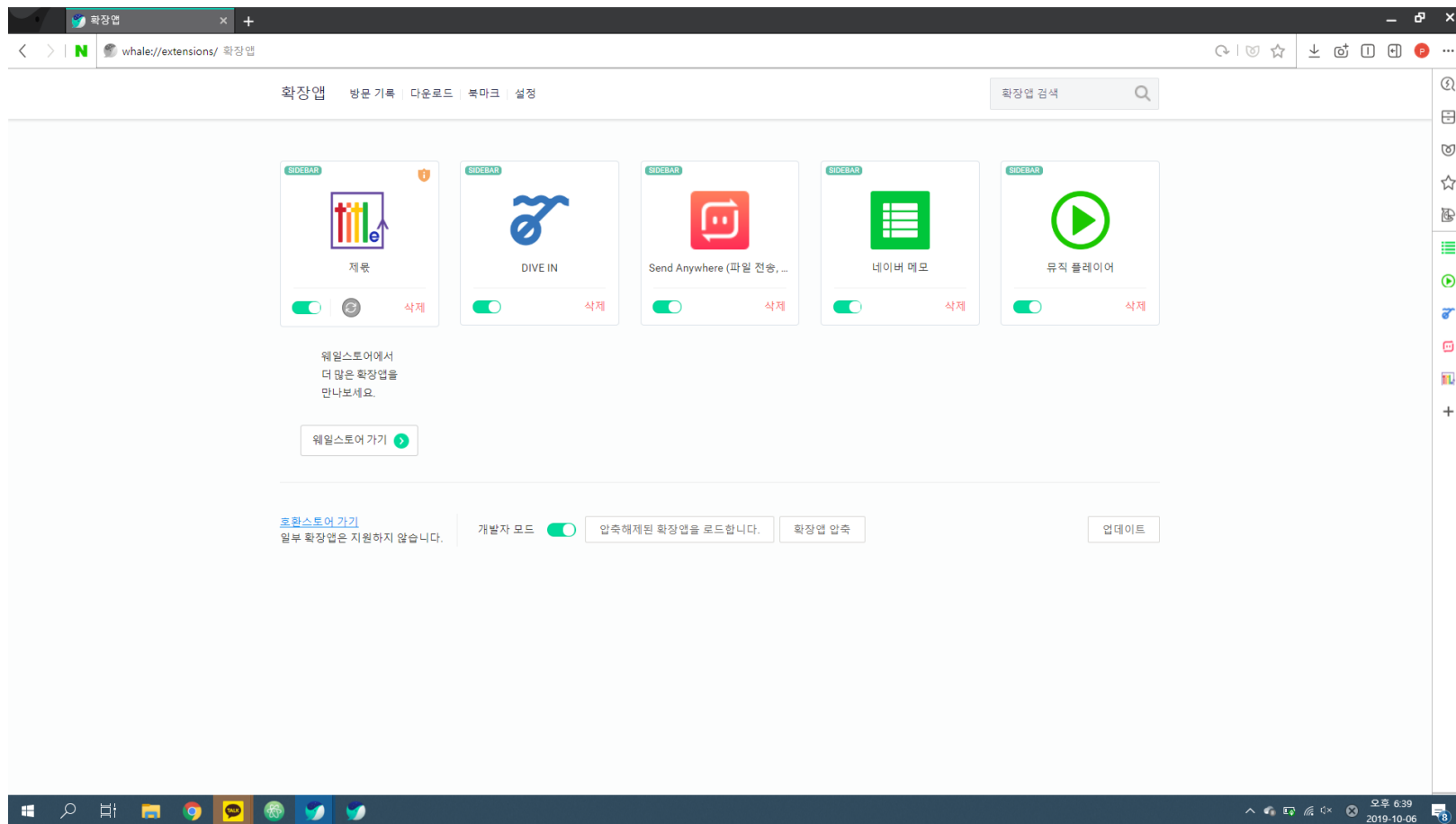


--- 3.9389870166778564 seconds ---

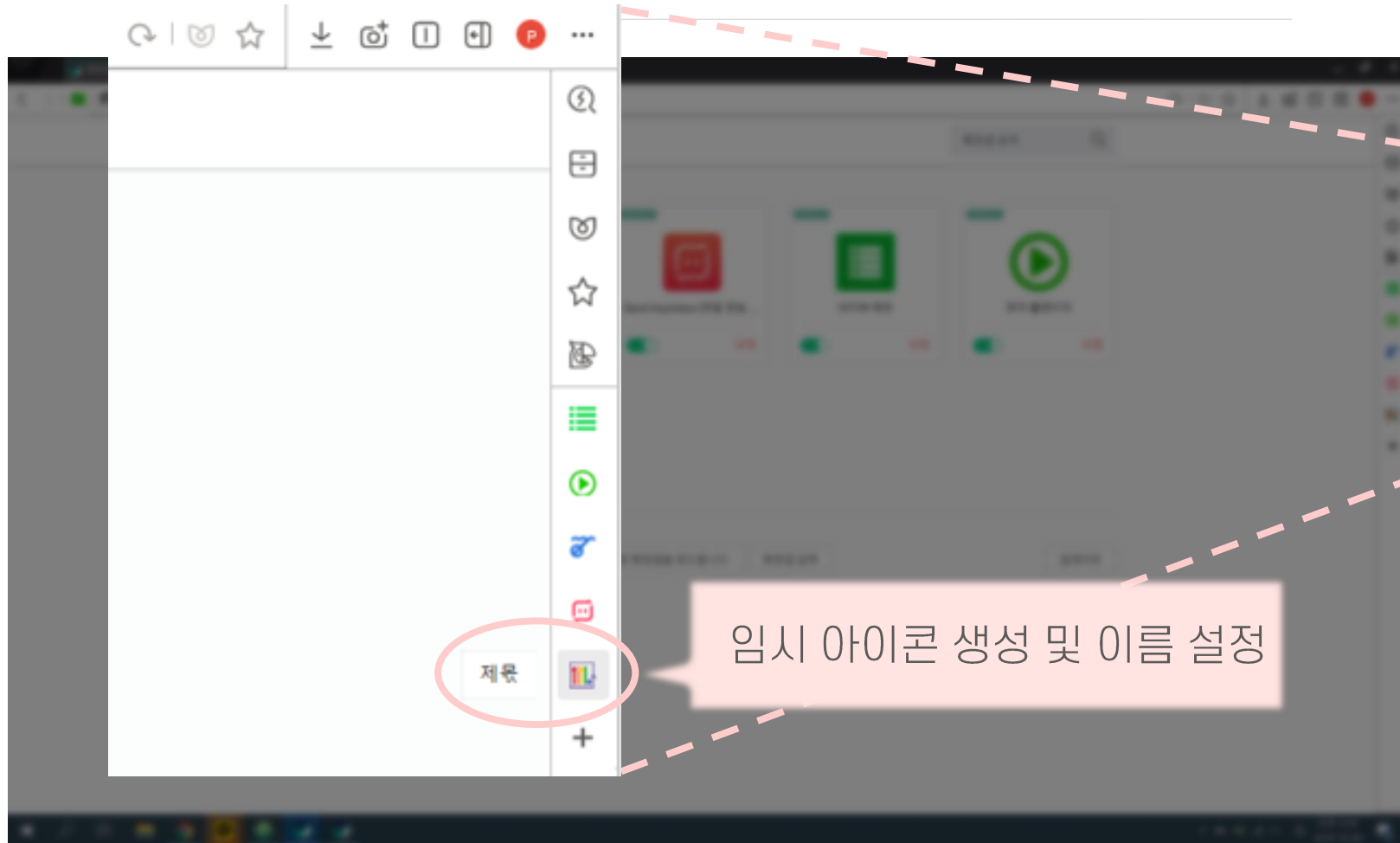
Multiprocessing 모듈 사용 시
대략 4초로 속도향상을 증명



아이콘 생성



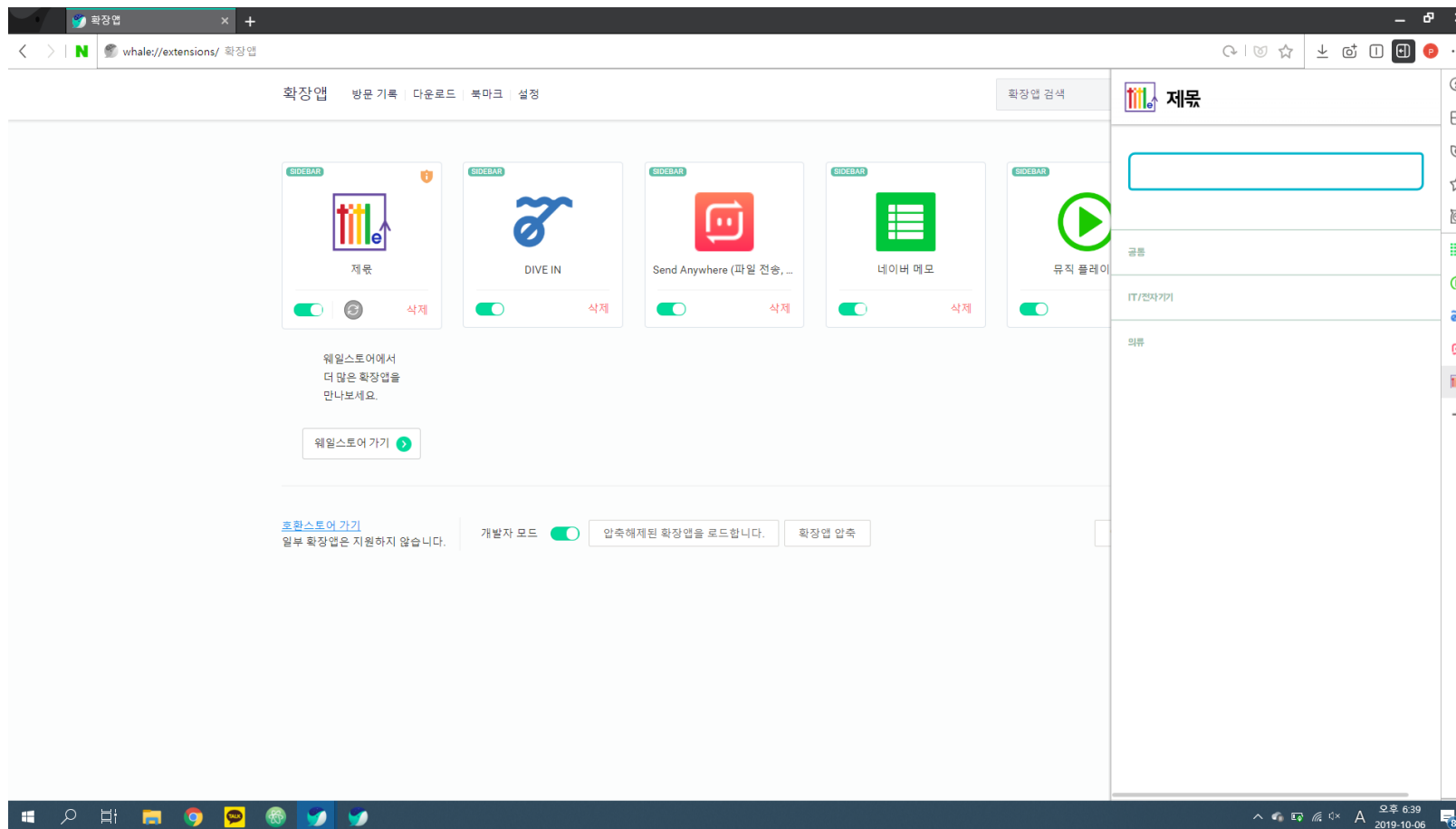
아이콘 생성



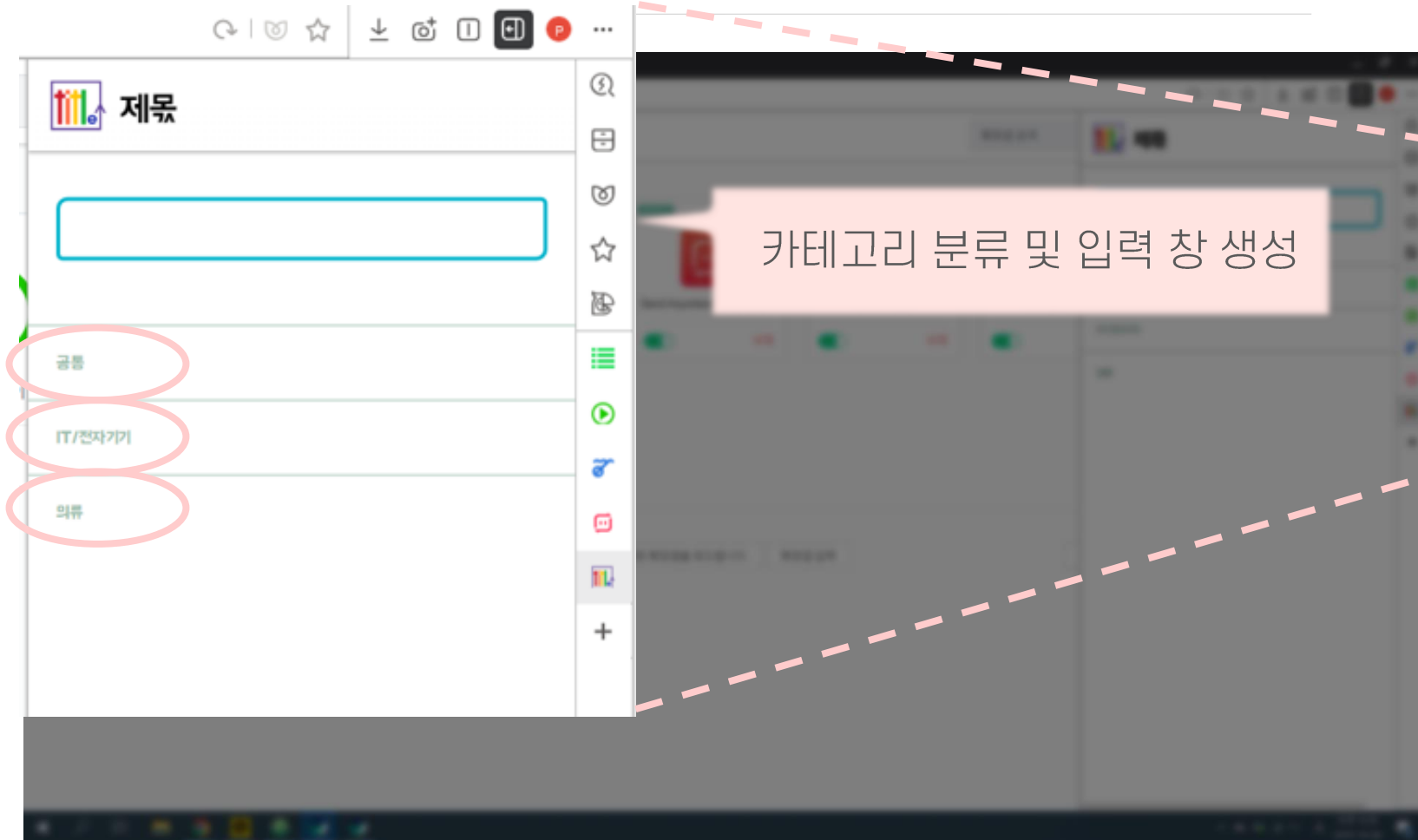
임시 아이콘 생성 및 이름 설정



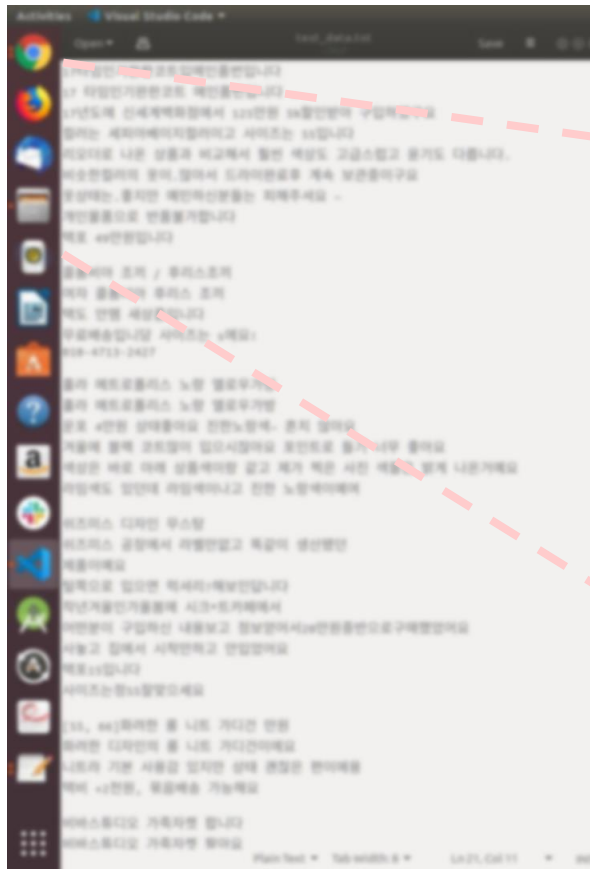
아이콘 생성



아이콘 생성



임시 데이터 생성



17타임인기완판코트입메인품번입니다
17 타임인기완판코트 메인품번입니다 .
17년도에 신세계백화점에서 125만원 5%할인받아 구입하였구요
컬러는 세피아베이지컬러이고 사이즈는 55입니다
리오더로 나온 상품과 비교해서 훨씬 색상도 고급스럽고 윤기도 다릅니다.
비슷한컬러의 옷이 많아서 드라이완료후 계속 보관중이구요
옷상태는 좋지만 예민하신분들은 피해주세요 ~
개인물품으로 반품불가합니다
택포 49만원입니다

txt 형태로 임시 데이터 추출

데이터 토큰화

토큰화 및 품사 추출

문서 벡터화

```
'''
[['아 /Exclamation',
 '더빙 /Noun',
 '...' /Punctuation',
 '진짜 /Noun',
 '짜증나다 /Adjective',
 '독소리 /Noun'],
 '0']
2159921
[nltk data] Downloading package stopwords to
[nltk data] /home/minheelee/nltk data...
[nltk data] Package stopwords is already up-to-date!
2159921
49895
[['.' /Punctuation', 67778),
 ('영화 /Noun', 50818),
 ('하다 /Verb', 41209),
 ('이 /Josa', 38540),
 ('보다 /Verb', 38538),
 ('의 /Josa', 30188),
 ('...' /Punctuation', 29055),
 ('가 /Josa', 26627),
 ('에 /Josa', 26468),
 ('을 /Josa', 23118)]
/home/minheelee/.local/lib/python3.6/site-packages/gensim/
warnings.warn("The parameter 'iter' is deprecated, will
/home/minheelee/.local/lib/python3.6/site-packages/gensim/
warnings.warn("The parameter 'size' is deprecated, will
[['공포 /Noun', 0.3906925320625305),
 ('코메디 /Noun', 0.38054558634757996),
 ('형사물 /Noun', 0.37719589471817017),
 ('슬래셔 /Noun', 0.37646836042404175),
 ('공포물 /Noun', 0.3637131452560425),
 ('로드무비 /Noun', 0.36252540349960327),
 ('호러물 /Noun', 0.3546419143676758),
 ('공포영화 /Noun', 0.3487427830696106),
 ('느와르 /Noun', 0.34804752469062805),
 ('카지노 /Noun', 0.3411845564842224)]
[['ㅋㅋ /KoreanParticle', 0.3027341961860657),
 ('ㅋ /KoreanParticle', 0.26884764432907104),
 ('SNL /Alpha', 0.2465825080871582),
 ('코미 /Noun', 0.2433760017156601),
 ('ㅋㅋㅋ /KoreanParticle', 0.2379237562417984),
 ('^^^ /Punctuation', 0.2288137972354889),
 ('...' /Punctuation', 0.2245882749557495),
 ('드라이버 /Noun', 0.21911387145519257),
 ('^^;; /Punctuation', 0.21901234984397888),
 ('>< /Punctuation', 0.21732869744300842)]
```

```
[(['.' /Punctuation', 67778),
 ('영화 /Noun', 50818),
 ('하다 /Verb', 41209),
 ('이 /Josa', 38540),
 ('보다 /Verb', 38538),
 ('의 /Josa', 30188),
 ('...' /Punctuation', 29055),
 ('가 /Josa', 26627),
 ('에 /Josa', 26468),
 ('을 /Josa', 23118)]
```

빈도수 순 출력
문장 구성에 따른 비슷한 토큰 출력

```
[('공포 /Noun', 0.3906925320625305),
 ('코메디 /Noun', 0.38054558634757996),
 ('형사물 /Noun', 0.37719589471817017),
 ('슬래셔 /Noun', 0.37646836042404175),
 ('공포물 /Noun', 0.3637131452560425),
 ('로드무비 /Noun', 0.36252540349960327),
```

자연어처리 API

KoNLPy

데이터 전처리
형태소 분석 품사 태깅

NLTK

데이터 탐색 요약
(토큰 수, 빈도 수 등)

Gensim

문서 벡터 추출
여러 문장에 내재되어 있는 규칙 토픽 분석

다음주 개발 계획

박시현

데이터 크롤링 및 데이터셋 구축
데이터 전처리 학습

박진영

웹 프론트엔드
확장 프로그램 발전

이민희

크롤링한 데이터 문장 단위로 쪼개기 및 API 추가 학습
미리 정의된 사전에 중고거래용 단어 추가



감사합니다!

