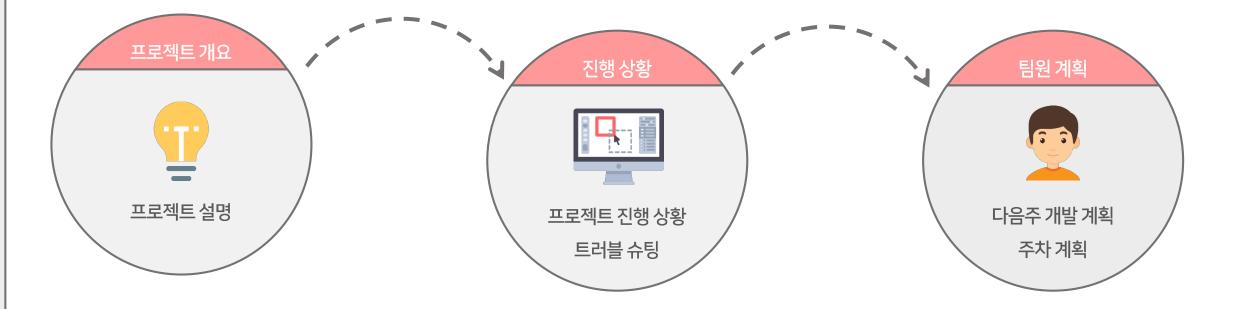


팀명 : 제몫 박시현 + 박진영 + 이민희













## 프로젝트 설명

♥무료배송♥ 워싱 면 라이더 자켓 판매 ◎ ◎

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

서버 변경 및 환경설정

서버 단 두 프로그램 연동

박진영

변경 전 제목 마우스오버시 툴팁 구현

프론트엔드 수정, 프론트엔드 상의 통신 구현

이민희

모델 파라미터 변경

상세 코드 구현



# AWS 인스턴스 생성

<b>탄력적 IP 주소</b> (1) 탄력적 IP 주소는 AWS 7	예정에 할당하고 인터넷에	서 접속할 수 있는 정적 퍼블릭 IPv4 주소?	입니다. 자세히 알아보기 🔼
Q 퍼블릭 <b>IPv4</b> 주소 <b>:</b> 15.1	164.195.30 ×	필터 지우기	
Name	퍼블릭 IPv4 주소	할당 ID	연결된 인스턴스
✓	15.164.195.30	eipalloc-08a335b8f3ab94cb6	i-0ffa6b048849cd6da 🛂

CPU 4개 인스턴트 유형: c5d.xlarge

키페어를 통해 서버에 접근

#### 서버 단 프로그램 연결

크롤링 후 자연어처리 함수를 통해 크롤링 한 데이터 값 (제목, 가격, 본문을 리스트로) 전달 키워드에 해당하는 값을 출력 (리스트)

뒤에서 결과값 설명 예정

## 서버의 크롬 브라우저 실행 오류

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

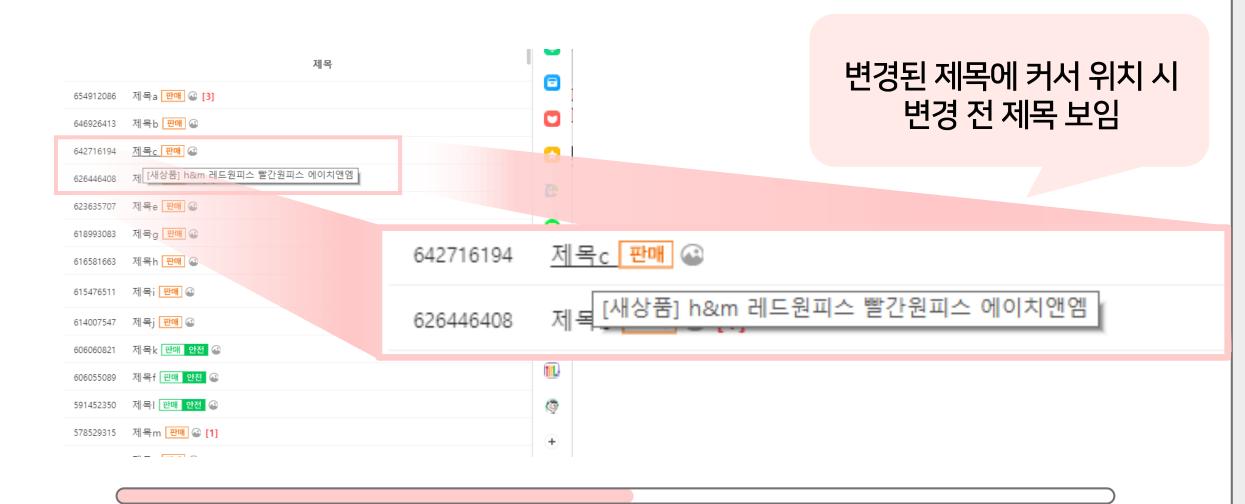
데이터 크롤링 시 크롬 브라우저를 여는 과정에서 발생하는 오류 크롬 브라우저 실행 시 띄워지는 웹 복구 창의 문제라고 생각



리눅스 가상환경을 사용

# 제목 변경 시 툴팁 생성

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



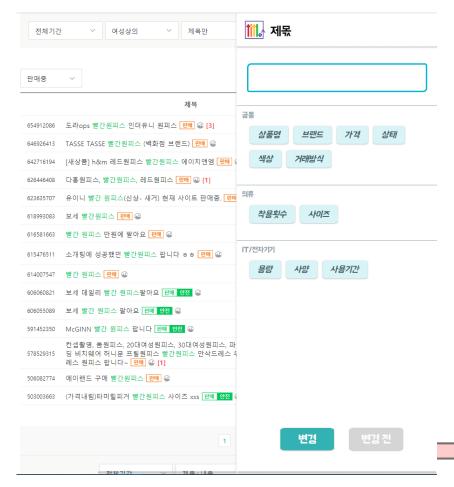




## 테스트 및 프론트엔드 수정

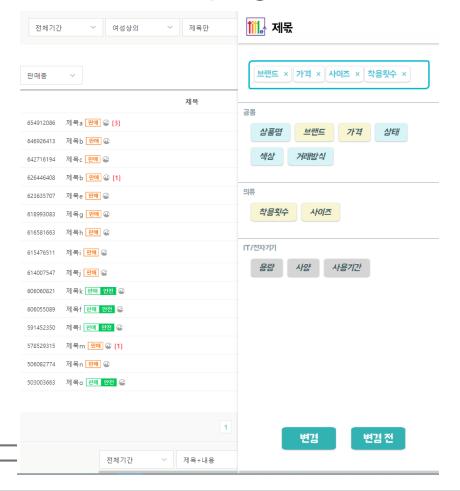
• • • • • • • • • • • • • • •

#### 변경된 제목이 없을 때



키워드 추가 & 변경 버튼 클릭

#### 제목 변경 후





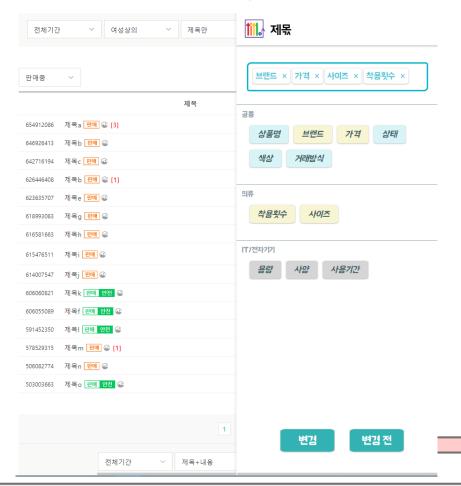




# 테스트 및 프론트엔드 수정

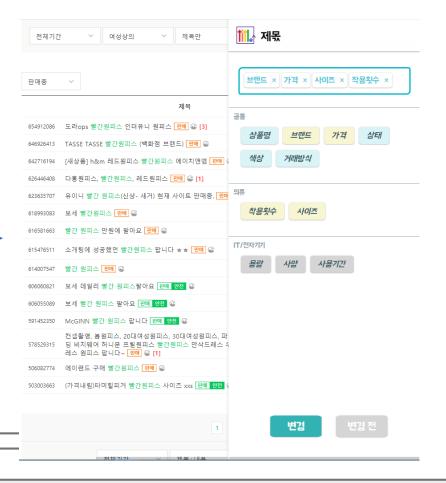
• • • • • • • • • • • • • • •

#### 제목 변경 후



#### 변경 전 버튼 클릭

#### 제목 변경 후







since 2003.12.10.

♣ 17,942,555 a ★ 즐겨찾는 멤버

☑ 게시판 구독수

▣ 전체글보기

᠍ 공지사항

- ᠍ 이용정책

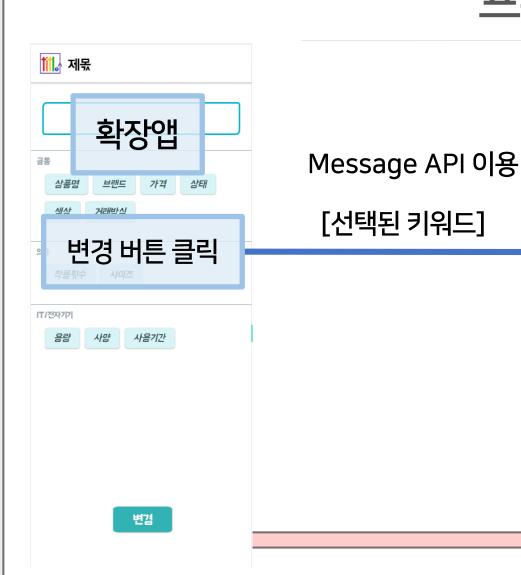
중고나라★새소식

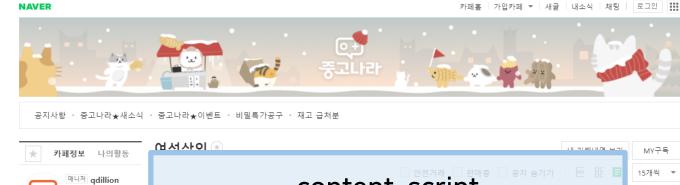
᠍ 중고나라★이벤트

- ᠍ 당첨자발표/종료

카페 가입하기

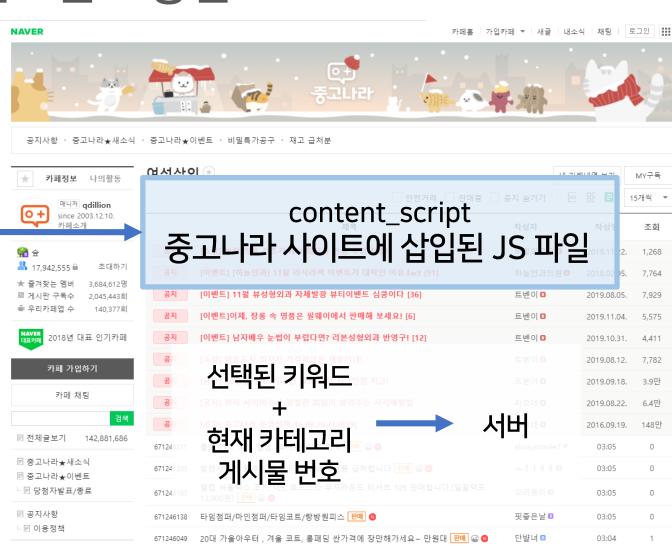
카페 채팅











since 2003.12.10.

<sup>♣</sup> 17,942,555 a ★ 즐겨찾는 멤버

᠍ 게시판 구독수

카페 가입하기

카페 채팅

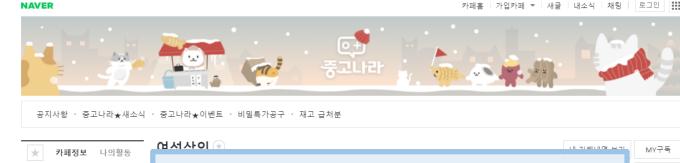
중고나라★새소식

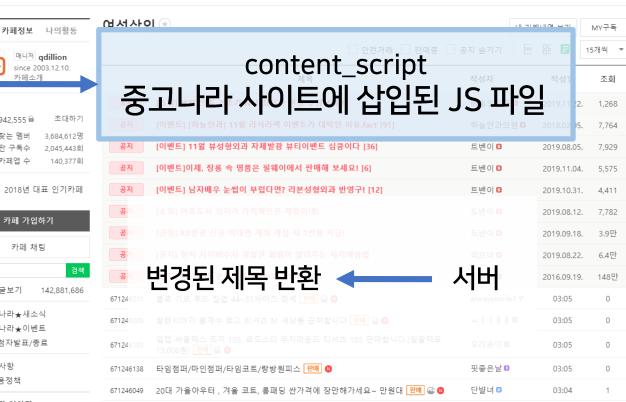
᠍ 중고나라★이벤트 └悥 당첨자발표/종료

᠍ 공지사항

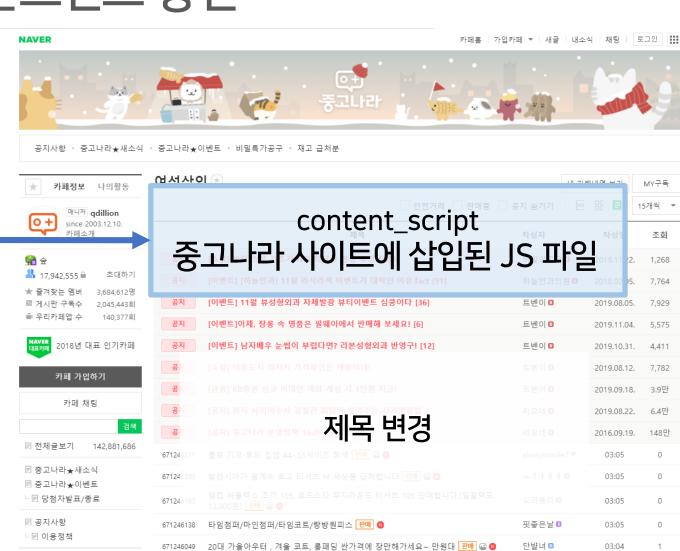
- ᠍ 이용정책

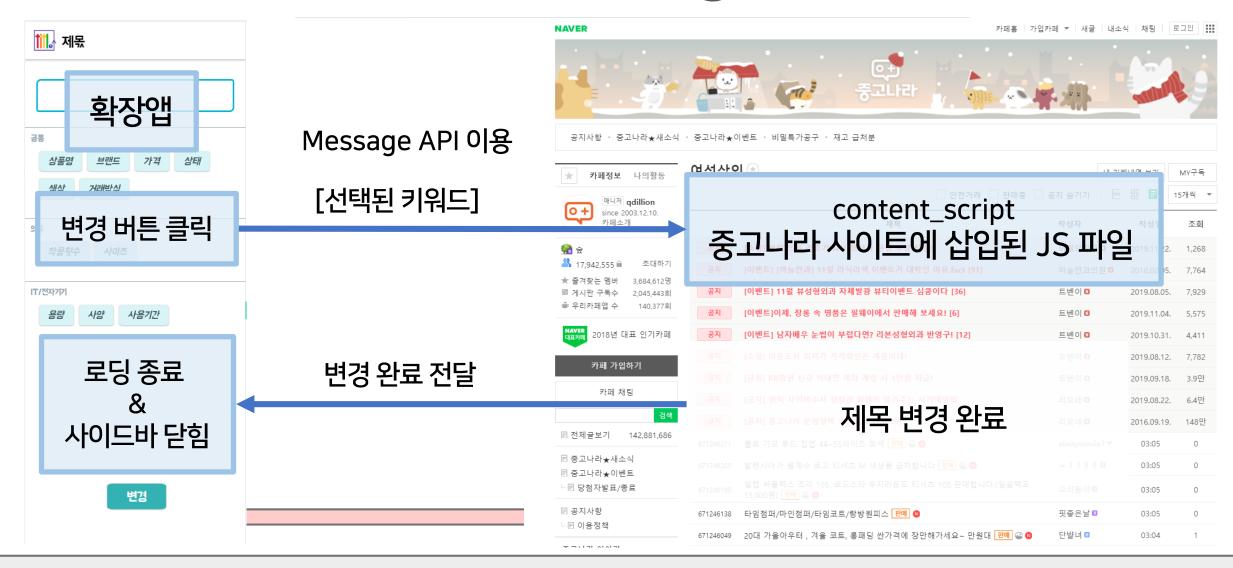












## 사전 추출 어려움

특징 (재질/무늬)



옷 카테고리

이름

Ex)

체크 린넨 스트라이프 꽈배기 자켓

셔츠

후드 집업

# 신발 이름 다양성

[새상품] 아디다스 알파바운스 슬리퍼 검흰 판매 🔐 🚯	조던11 브레드 270사이즈 판매합니다. 판매 🚳 🔕
나이키 에어포스1 고어텍스 로우 옐로우 255사이즈 나코틴	[255] 나이키 데이브레이크 파금 판매 🚳 🕦
나이키 피스마이너스원 파라노이즈 255 검흰 판매 🐼 🔕	265. 나이키 피스마이너스원 에어포스 1 파라노이즈 구매합니다. 최고가 🕝 🕦
오프화이트 Off-White 슬리퍼 250사이즈 새제품 판매합니	[290] 에어맥스 97 나이키 런던 220,000원 판매 ຝ №
트래비스 스캇 포스 삽니다. 255 260 265 270 275 판매 🔐	피마원 에어포스 270-275 삽니다 실착용 판매 ຝ 0 0
아디다스 오즈위고 흰파형 250 (ee7009) 팝니다. 판매 🔐 🕻	뉴발란스 999시리즈ML 999 WEU 상태 초S급(270) 판매 ຝ ₫
y3 y-3 카이와 검흰 새제품 250사이즈 팝니다. 판매 ὧ №	(270,275)나이키 피마원 에어포스1 검흰 삽니다. 판매 ◎
(235)조던1 백보드3.0 GS <mark>판매 </mark>	

#### 이름 처리 변경

• • • • • • • • • • • • • • •

특징 + 카테고리

옷 이름

Ex)

스트라이프 가죽 코트 면 팬츠 집업 etc 연달아 나오는 명사, 알파벳, 숫자 최대 3개 조합을 찾고 그 중 가장 긴 조합 반환

신발 이름

#### 신발 이름 추출

```
for t, p in all docs list:
   if p == 'Noun' or p == 'Alpha' or p == 'Number':
       if t not in ["사이즈", "size", "싸이즈", "싸쥬", "mm"]:
           if t in ["제품", "정품", "상태", "사용감", "가격", "판매가", "택포", "착불", "희망가", "운포"]:
              noun appear = 0
              count = count + 1
              if count > 3:
                  break
              noun lists.append(noun list)
              noun list = []
              continue
           noun appear = 1
           pre_p = p
                                                                                   elif noun_appear != 0:
           noun list.append(t)
       else:
                                                                                        noun appear = 0
           if pre_p == 'Number':
                                                                                        count = count + 1
              noun list.pop()
                                                                                        if count > 3:
              noun appear = 0
                                                                                            break
              count = count + 1
              if count > 3:
                                                                                        noun lists.append(noun list)
                  break
                                                                                        noun list = []
              noun lists.append(noun list)
              noun list = []
                                                                               return name = max(noun lists, key=len)
              continue
```

#### 이론과 실전의 차이



#### 이론과 실전의 차이

```
['나이키 문레이서 새상품 290사이즈 운포 13.9', 800,'새상품입니다 지난주 직구한 후 시착한번 안하고 판매합니다'], ['닥터마틴 Joyce 260mm 판매합니다.', 550,'닥터마틴 Joyce 260mm 판매합니다. 시착만한 새제품입니다 사이즈미스로 판매합['(새제품)아디다스 알파바운스 블랙(B41720) 255 판매합니다.', 110,' (새제품)아디다스 알파바운스 블랙(B41720) 255사이즈 ['아식스 x 레이닝챔프 젤카야노25 사이즈:265', 320,'상품: 아식스 x 레이닝챔프 젤카야노 25 사이즈: 265 상태: 새상품 (한['[M 285/ W 290] 나이키X오프화이트 줌테라 키어거5 그린 사이즈 : W290 / I
```

```
['나이키', '나이키', '13.9', '?', '
['닥터마틴 Joyce', '?', '12만원', '
['아디다스 알파 바운스 블랙', '아디니
['아식스 x 레이 닝 챔프 젤', '아식스
['나이키 x 오프 화이트', '나이키', (env)
```

#### 인생은 내 마음 같지 않음

```
['나이키 문레이서 새상품 290사이즈 운포 13.9', 800,'새상품입니다 지 ∨ all_docs_list: [('나이키', 'Noun'), ('문', 'Modi
['닥터마틴 Joyce 260mm 판매합니다.', 550,'닥터마틴 Joyce 260mm 판
                                             > 00: ('나이키', 'Noun')
['(새제품)아디다스 알파바운스 블랙(B41720) 255 판매합니다.', 110,
                                            > 01: ('문', 'Modifier')
['아식스 x 레이닝챔프 젤카야노25 사이즈:265', 320,'상품: 아식스 x
['[M 285/ W 290] 나이키X오프화이트 줌테라 키어거5 그린', 190, '나(
                                             > 02: ('레이서', 'Noun')
         ['나이키', '나이키', '13.9', '?', '
                                            |> 03: ('새', 'Modifier')
         ['닥터마틴 Joyce', '?', '12만원', '
                                            |> 04: ('상품', 'Noun')
         ['아디다스 알파 바운스 블랙', '아디
                                             > 05: ('290', 'Number')
         ['아식스 x 레이 닝 챔프 젤', '아식리
                                             > 06: ('사이즈', 'Noun')
         ['나이키 x 오프 화이트', '나이<u>키'</u>
                                       ∨ noun_lists: [['나이키'], ['레이서'], ['상품']]
         (env)
                                        > 0: ['나이키']
                                        > 1: ['레이서']
                                        > 2: ['상품']
```

## 그럭저럭 선방

```
['나이키 문레이서 새상품 290사이즈 운포 13.9', 800,'새상품입니다 지난주|
['닥터마틴 Joyce 260mm 판매합니다.', 550,'닥터마틴 Joyce 260mm 판매합니<mark>V noun_lists: [['닥터마틴', 'Joyce</mark>
['(새제품)아디다스 알파바운스 블랙(B41720) 255 판매합니다.', 110,' (새저
                                            > 0: ['닥터마틴', 'Joyce']
['아식스 x 레이닝챔프 젤카야노25 사이즈:265', 320,'상품: 아식스 x 레이닝
['[M 285/ W 290] 나이키X오프화이트 줌테라 키어거5 그린', 190, '나이키XS
                                            > 1: ['판매']
        ['나이키', '나이키', '13.9', '?', '
                                            > 2: ['닥터마틴', 'Joyce']
        ['닥터마틴 Joyce', '?', '12만원', '
        ['아디다스 알파 바운스 블랙', '아디!
        ['아식스 x 레이 닝 챔프 젤',
                                     ∨ noun_lists: [['제품'], ['아디다스', '알파'
        ['나이키 x 오프 화이트', '나이
                                      > 0: ['제품']
        (env)
                                      > 1: ['아디다스', '알파', '바운스', '블랙']
                                      > 2: ['B', '41720']
```

['나이키 문레이서 새상품 290사이즈 운포 13.9', 800,'새상품입니다 7

#### 의도치 않은 토큰화

```
['닥터마틴 Joyce 260mm 판매합니다.', 550,'닥터마틴 Joyce 260mm 판 ['(새제품)아디다스 알파바운스 블랙(B41720) 255 판매합니다.', 110,'
['아식스 x 레이닝챔프 젤카야노25 사이즈:265', 320,'상품: 아식스 x
['[M 285/ W 290] 나이키X오프화이트 줌테라 키어거5 그린', 190, '나이 ['나이키', '나이키', '13.9', ''.', ' 002: ('레이' 04: ('챔프' 06: ('구터마틴 Joyce', '?', '12만원', ' 003: ('본', 06: ('캠프' 06: ('구나이키 X 오프 화이트', '나이키', '아니 005: ('프다' 006: ('크다' 007: ('건5', 008: ('사이: '008: ('
```

```
∨ all_docs_list: [('아식스', 'Noun'),
 > 000: ('아식스', 'Noun')
 > 001: ('x', > 03: ('/', 'Punctuation') > 002: ('레이' > 04: ('W', 'Alpha')
 > 003: ('닝', > 05: ('290', 'Number')
 > 004: ('챔프' > 06: (']', 'Punctuation')
 > 005: ('젤', > 07: ('나이키', 'Noun')
 > 006: ('크다' > 08: ('X', 'Alpha')
                 > 09: ('오프', 'Noun')
                 > 10: ('화이트', 'Noun')
                 > 11: ('주다', 'Verb')
 > 009: (':',
                 > 12: ('키', 'Noun')
 > 010: ('265' |> 13: ('어', 'Suffix')
                 > 14: ('거', 'Noun')
```

#### Seed 작업

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

베이지 체크 자켓 블루 기모 후드 티

```
for word, prob in sorted(pred_pos.items(), key=lambda x:-x[1])[:500]:
    if word.split('/')[1] != 'Noun':
        continue
    if word in seed_words:
        continue
    if word in real_nseed_words:
        continue
    idx = vocab_to_idx[word]
    count = idx_to_count[idx]
    f.write(str(word)+'\n')
    print(str(word))
```

#### 이번주 총 결과 화면

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

```
', '뉴발', 45000, '?', '?',
     ', '노스페이스', 9000, '?',
     ', '뉴발란스', 150000, '검정
'투맨 ', '스파이더', 120000, '검정
     ', '?', 5000, '?', '?', '5천
'패딩 조끼 ', '?', 100000, '?', '틱
'라이더 자켓 ', '카파', 10000, '?'
 '코트 ', '?', 139000, '60네이비색'
 '롱패딩 ', '코오롱', 11000, '?', '
 '경량 패딩 ', '경량', 40000, '블랙
''후드 점퍼 ', '아베크롬비', '?', '
'셔츠 ', '?', 8000, '?', '택배', '
'니트 ', '?', 65000, '?', '?', 100
['후드 후디 ', '나이키', 120000, '?
['패딩 ', '나이키', '?', '?', '?',
(example)
```

```
['니트 스웨터 ', '?', 10000, '?', '택배', '?']
['롱패딩 ', '뉴발란스', '200,000', '?', '?', '?']
['자켓 ', '?', 240000, '?', '택배', 'XXL']
['후드 니트 가디건 ', '라코스테', '?', '?', '?', 'LACOSTE']
['코트 캐시미어 히든 그레이 싱글 코트 ', '?', 80000, '?', '?', 100]
['라이더 자켓 ', '꼼데가르송', 180000, '?', '?', '?']
['', '?', '?', '?', '?', '?']
['빅크록 투맨 ', '라코스테', 100000, '?', '?', '?']
['트레이닝 져지 상의 ', '나이키', 15000, '?', '?', '?']
['후드 티 ', '?', 130000, '?', '?', '?']
['폴햄 베이지 코트 폴햄 코트 베이지 ', '폴햄', 60000, '?', '?', 'L']
['점퍼 ', '?', 15000, '?', '?', '?']
['에센셜 숏패딩 베이지 ', '?', 139000, '?', '?', 's']
['사파리 바람막이 윈드 자켓 ', '컬럼비아', 40000, '?', '?', 'XL']
['엠폴햄 스트라이프 코트 ', '엠폴햄', 70000, '?', '?', '?']
(example)
```

#### 이번주 총 결과 화면

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

```
니고~~ , : , ユಶ・ಶ라건 , : , 곡애포국기네 ,
['컨 스타 바이지', '?', '69,000원', '?', '택배', 270]
['남자 운동화 포스 하이', '나이키', '?', '?', '택배+직거래', 255]
['조던 260', '나이키', '?', '?', '택배+직거래', '?']
['윈드 79 파', '나이키', '17', '?', '택배', '?']
['부스트 350 V 2 트리플 화이트 265', '아디다스', '?', '?', '택배+직거래', 265]
['에어 맥스 95 모와브 275', '나이키', '5만원', '?', '택배', '?']
['원 본디 5', '귀하다', '원', '?', '?', 275]
['피스 마이너스', '에어포스', '?', '?', '?', '275mm']
['Paracia', '?', '158,000원', '레드색', '택배', 270]
['파라 노이즈 275 검빨', '나이키', '?', '?', '직거래', 275]
['사카이 블 레이져 노빨 260 나코탭', '나이키', '?', '?', '?', '?']
(example)
```

#### 모델 사전 작업 항시 작업 중^^

• • • • • • • • • • • • • •

```
≡ m_brand_name_ner_30_4.txt
```

- = m\_brand\_name\_ner\_80\_4(1).txt
- = m\_clothes\_model40000\_30\_4
- m\_clothes\_model40000\_80\_4
- = m\_clothes\_name\_ner\_30\_4(2).txt
- = m\_clothes\_name\_ner\_80\_4.txt
- = m\_clothes\_name\_ner\_80\_4(1).txt

```
phone_model30000_80_8
```

```
≡ it_model20000_40_4
```

it\_model20000\_50\_8

it\_model20000\_80\_10

여자.. 남자도 나누고.. 옷.. 신발도.. 나눠서 모델 돌려.. Ner도 각각.. 만들어줘..



#### 모델 사전 작업 항시 작업 중^^

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

```
> word2vec.wv.most_similar('아이폰/Noun', topn=10): [('담요/Noun'...
> 0: ('담요/Noun', 0.40604114532470703)
> 1: ('아동복/Noun', 0.38518714904785156)
> 2: ('즉석식품/Noun', 0.3788599371910095)
> 3: ('멜리사/Noun', 0.3555731177330017)
> 4: ('방진/Noun', 0.3474213182926178)
> 5: ('앤더/Noun', 0.3381642997264862)
> 6: ('홍대/Noun', 0.32922106981277466)
> 7: ('드라잉/Noun', 0.3215567469596863)
> 8: ('석화/Noun', 0.32143378257751465)
> 9: ('닌텐도위/Noun', 0.3195466995239258)
__len__: 10
```

```
> word2vec.wv.most_similar('아이폰/Noun', topn=10): [('샌
> 0: ('샌디스크/Noun', 0.37357914447784424)
> 1: ('아동복/Noun', 0.3426370620727539)
> 2: ('방진/Noun', 0.33846020698547363)
> 3: ('나이스/Noun', 0.3319269120693207)
> 4: ('짤/Noun', 0.32969167828559875)
> 5: ('아이패드미니/Noun', 0.32353514432907104)
> 6: ('가차/Noun', 0.31763720512390137)
> 7: ('음료수/Noun', 0.31325504183769226)
> 8: ('멜리사/Noun', 0.301723450422287)
> 9: ('렌즈/Noun', 0.2977134883403778)
__len__: 10
```

It 기기 의미적으로 비슷한 게 필요해.. Window 크기 늘려..

# 다음주 개발 계획

박시현

서버 단 통신 모듈 구현

단위 테스트 및 최종 데모 준비

박진영

프론트엔드 단 통신 모듈 구현

단위 테스트 및 최종 데모 준비

이민희

모델 파라미터 변경

성능 및 정확도 향상

# 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															
크롤링ㆍ전처리															
데이터셋 구축								중간 고사							최종
서버 구축								& 중간						최종 데모	최종 리포 트
통신 모듈 작업								데모 준비							제출
서버 통신 구현															
데모 준비 · 서비스 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	최종 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 리포
텍스트 대치 구현								& 중간 데모						데모	트 제출
자연어 처리 모델 구축								준비							"=
웹 페이지 쿠키 구현															
모듈 연결															
유닛 · 전체 테스트															



