

MUSINSA

Recommending Items & Predicting star rating

PROJECT



TIMELINE

PROJECT ISSUE

- ◆ ISSUE POINT
- ◆ CHALLENGE WAY

PROJECT DETAIL

- ◆ RECOMMENDING ITEMS
- ◆ PREDICTING REVIEW STAR RATE
- ◆ TEAM MEMBER & ROLE

APPENDIX

TIMELINE

06

1_WEEK

프로젝트 주제 탐색

- 브레인스토밍
- 주제 선정

데이터 셋 구축

- 크롤링 및 데이터 핸들링

2_WEEK

토픽 모델링 / 시각화

ML [텍스트마이닝을 통한 상품 추천]

- 모델 성능 개선 연구

3_WEEK

ML [머신러닝을 통한 감성분석]

- 사전 생성 [라벨링]

DL [딥러닝을 통한 별점 예측]

- 래퍼런스 논문 리뷰

06

4_WEEK

DL [딥러닝을 통한 별점 예측]

- 리뷰 데이터셋 증축
- 래퍼런스 코드 커스텀

5_WEEK

모델 성능 검증 및 결과물 도출

- 데이터셋 증축 및 성능 보완

LET'S THINK ABOUT IT
IN TERMS OF MUSINSA

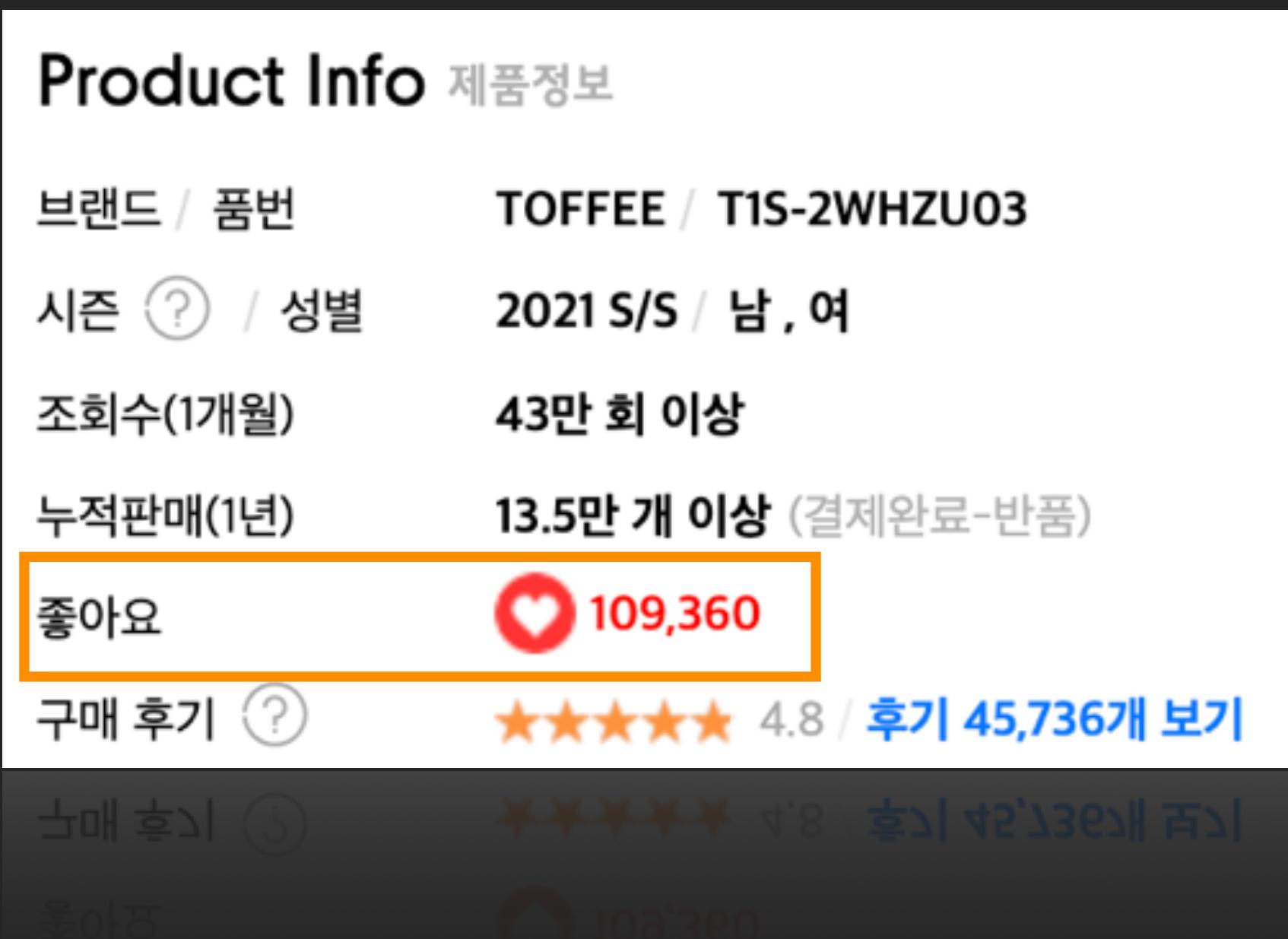
PROJECT DETAIL

RECOMMENDING ITEMS

다음 해 매출 상위 상품 예측

ISSUE POINT

POST 방식



requests.get()

데이터 호출 목적

✓ requests.post()

데이터 저장 목적

ISSUE POINT

상품별 좋아요 수 크롤링을 위한 상품별 URL 맨 끝번호 필요 == **relationIds**

<https://www.musinsa.com/app/goods/1778404>

Product Info 제품정보

브랜드 / 품번	TOFFEE / T1S-2WHZU03
시즌 ? / 성별	2021 S/S / 남, 여
조회수(1개월)	43만 회 이상
누적판매(1년)	13.5만 개 이상 (결제완료-반품)
좋아요	109,360
구매 후기 ?	★★★★★ 4.8 / 후기 45,736개 보기
구매 후기	★★★★★ 4.8 / 후기 45,736개 보기
좋아요	

✓ Payload → relationIds

Payload

Request Payload

```

{
  "relationIds": [
    "1778408",
    "1778406",
    "1778401",
    ...
  ]
}
  
```

CHALLENGE WAY

무신사 크롤링 데이터를 이용하여 새로운 데이터 셋 구축

$$BMI = \frac{\text{체중 (kg)}}{\text{신장 (m)}^2}$$

< 18.5
저체중

18.5 - 24.9
표준

25 - 29.9
과체중

30 - 34.9
비만

35 <
고도비만

<정상범위 기준표>

$$\text{실제구매율 (\%)} = \frac{\text{PURCHASE}}{\text{PAGEVIEW}}$$

$$\text{할인률 (\%)} = \frac{\text{정가} - \text{할인 후 가격}}{\text{정가}} \times 100$$

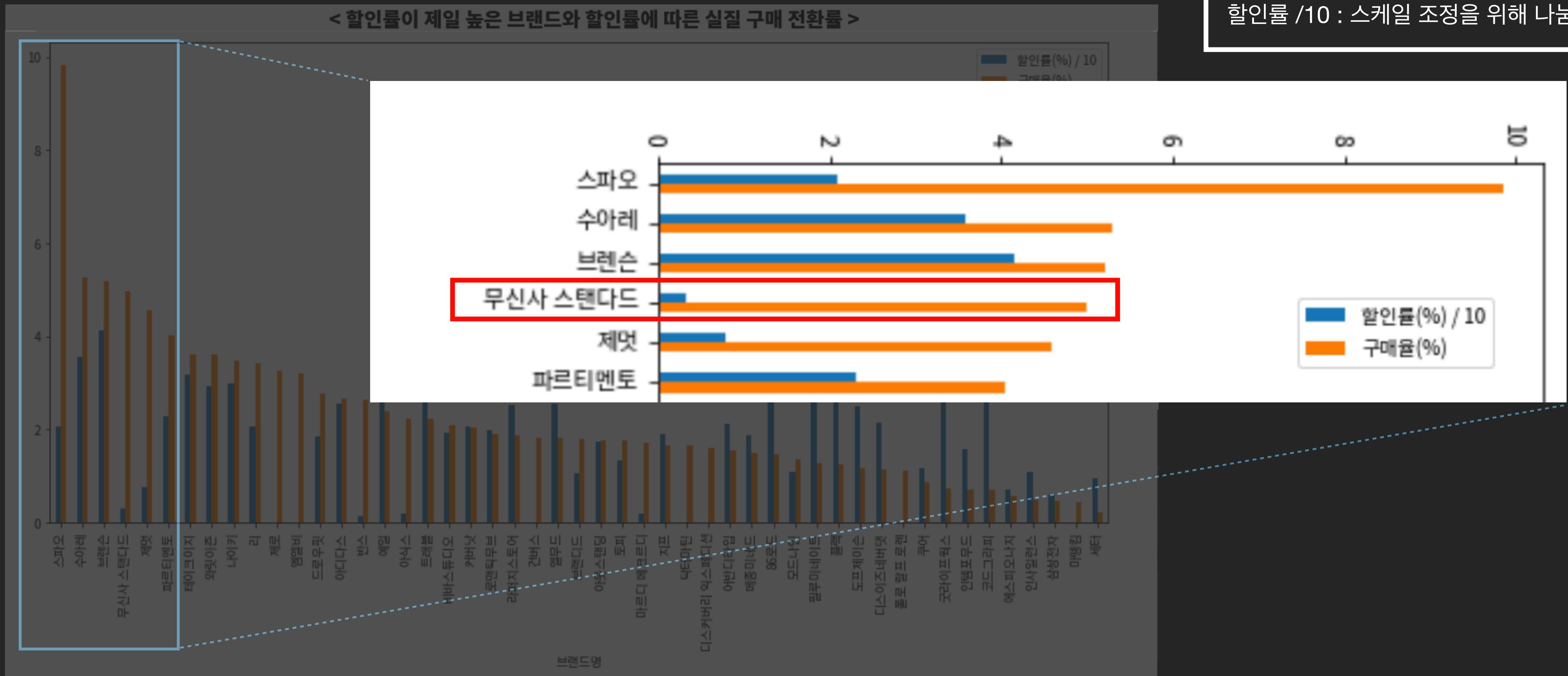
성별	키	몸무게	구매한 사이즈	대분류	BMI	정상범위
남성	182	72	XL	아우터	21.7	표준
남성	183	80	XL	아우터	23.9	과체중
남성	177	62	L	아우터	19.8	표준
남성	174	67	L	아우터	22.1	표준
남성	175	71	L	아우터	23.2	과체중

할인전가격	할인후가격	할인률(%)	pageview	purchase	실제구매율(%)
45000	39900	11.33	431174	134274	2.595124
39000	39000	0.00	390822	133499	2.846543
139000	139000	0.00	194820	42414	1.814239
95000	95000	0.00	144714	40383	2.325449
339000	225000	33.63	164277	10234	0.519143

DATA ANALYSIS RESULT

MUSINSA 사이트에서 MUSINSA STANDARD 의 랭킹 [구매 전환률 4위]

할인율에 비해 구매율이 다른 상위 브랜드들보다 상당히 높은 것을 확인할 수 있다. [실질적 1위]

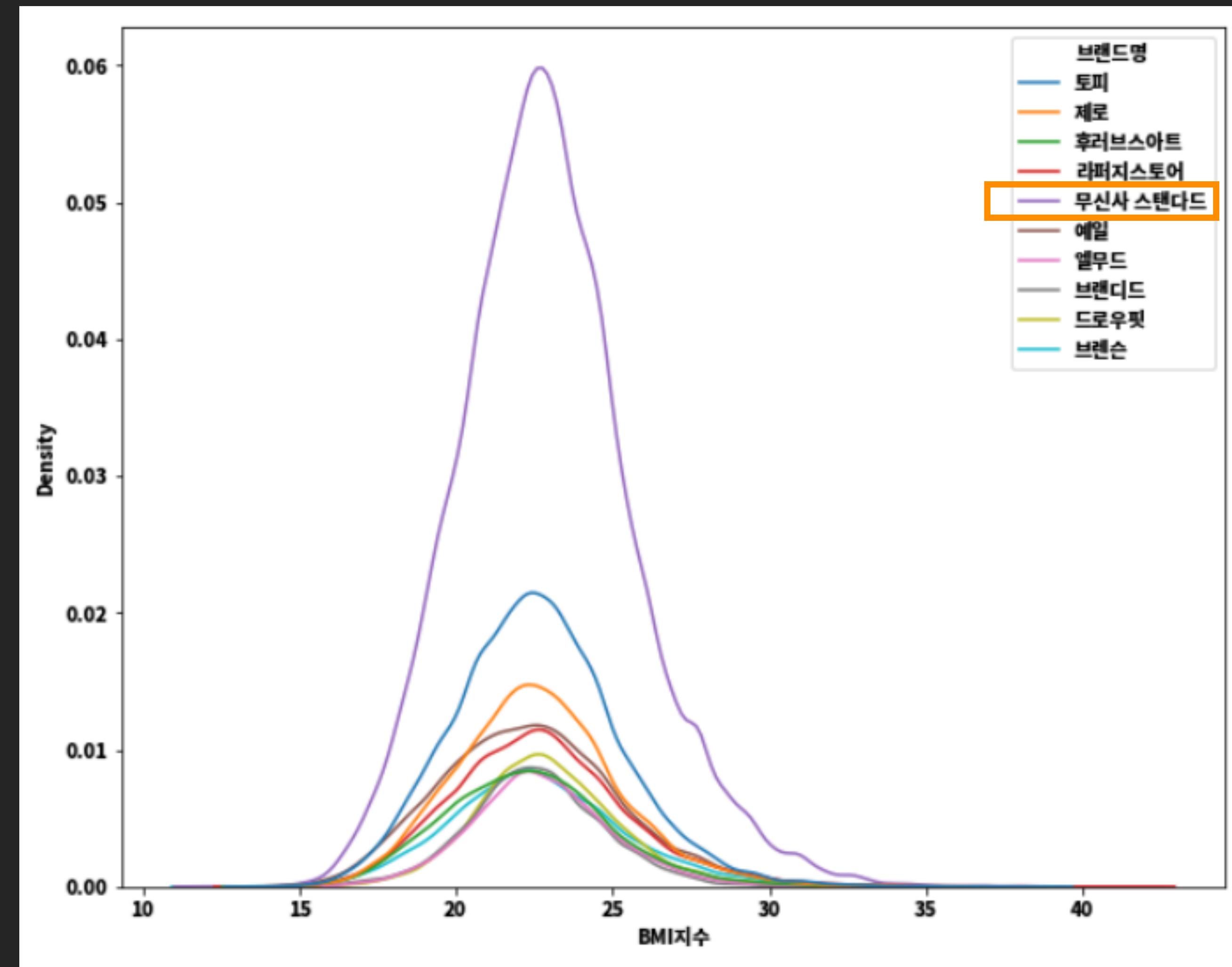


구매율 /12 : 조회수(1개월) / 누적판매수(1년)

할인률 /10 : 스케일 조정을 위해 나눔

DATA ANALYSIS RESULT

MUSINSA 사이트에서 MUSINSA STANDARD 의 랭킹



DATA ANALYSIS RESULT

무신사 > 랭킹 > 'TOP 990위' 까지 분석

카테고리 / 브랜드 부문으로 각각 순위 분석

Data Analysis Results	1위	2위	3위
전환률 (Category)	뷰티	모자	액세서리
조회수 (Category)	아우터	바지	선글라스 / 안경테
누적판매수 (Category)	선글라스 / 안경테	바지	액세서리
랭킹 990위 각 카테고리 수	상의	아우터	바지
할인률 1위 (Brand)	트래블	브렌슨	86로드
구매율 1위 (Brand)	스파오	수아레	브렌슨
3개월간 후기 수 기반 판매량 1위	트레이닝 / 조거팬츠	데님팬츠	슬랙스

Data Analysis Results	1위
주 구매고객 키 / BMI 치수	170 cm ~ 180cm / BMI 치수 : 20 ~ 25
BMI 기준 구매고객 브랜드 1위	무신사 스탠다드
좋아요 수 1위 (Brand)	커버낫 (Covernat) - 아우터

TARGET

대부분의 분야에서 상위 랭킹을 차지한 MUSINSA STANDARD

BMI 기준 구매고객 브랜드 1위	무신사 스탠다드
--------------------	----------

Brand List 브랜드 리스트

베스트브랜드 TOP100

1위 -

musinsa standard

무신사 스탠다드
musinsastandard

musinsa standard

무신사가 제안하는 새로운 기준

musinsa standard(무신사 스탠다드)는 대한민국을 대표하는 온라인 패션 스토어 무신사가 전개하는 자체상표(Private Brand)이며, 모던 베이식 캐주얼웨어를 지향합니다.

단순한 유통구조의 효율적인 운영으로 좋은 품질의 제품을 합리적인 가격에 선보입니다.

가치소비 시대의 새로운 기준을 제시합니다.

Let's think about it in terms of MUSINSA

입점 브랜드 판매량 < 자체 브랜드의 성장성

✓ Recommending Items limited to MUSINSA STANDARD

<무신사 스탠다드 상품 추천 시스템>

TARGET

MAKING RECOMMEND SYSTEM <‘MUSINSA STANDARD’ 상품 상세 설명과 큐레이팅 글과 유사도를 비교분석>

CURATING SHOP (2021 기준)

MUSINSA 의 패션 에디터 (전문가)들이 작성한 글.

주로 입점한 다양한 브랜드들에 대한 상품 추천 글을 작성한다.

ITEM DETAILS [MUSINSA STANDARD ONLY]

MUSINSA STANDARD 의 상품에 대한 설명과 특징 기술

전문 지식이 가미된 단어들과 내용이 주를 이룬다.

“ IN THIS PROJECT.. ”

1. 2021년도 CURATING SHOP 글의 업로드 날짜 분석
2. 분기별로 나누어 2022년도 추천 상품 제시

EXPECT EFFECTIVENESS

1. 매년 다음 시즌 및 분기별 추천 상품들을 리스트업 해볼 수 있다.
2. 다음해 MUSINSA 매출에 영향을 줄 수 있는 상품 [브랜드] 선정 후, 프로모션 / 이벤트 등 다양한 매출 상승 방법을 미리 계획할 수 있을 것으로 기대된다.

VISUALIZATION

WORD CLOUD : CURATING SHOP [2021년도 기준]



ITEM DETAILS [MUSINSA STANDARD ONLY]



ITEM RECOMMENDATION

유사도 바탕 최종 상품 추천을 위한 아이템 추천

이미지 > 허드/블레이저 제작 (무신사 스텠다드)

무신사 단독 우먼즈 글렌 체크 오버사이즈 블레이저 [다크 베이지] WOMENS GLEN CHECK OVERSIZE BLAZER [DARK BEIGE]



무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

브랜드 / 품번 MUSINSA STANDARD / MWJJK003-DB
시즌 ⓘ / 성별 2020 F/W 여
조회수(1개월) 5.6천 회 이상
누적판매(1년) 700개 이상 (결제완료-반품)
좋아요 3,044
구매 후기 ⓘ ★★★★★ 4.8 / 후기 708개 보기
#제작 #체크블레이저 #글렌체크블레이저 #글렌체크자켓 #체크자켓
#글렌체크제작 #무신사스탠다드우먼즈 #우먼즈아우터 #가을이아우터

19~28세, 여성 에게 인기 많은 상품 ▾

Delivery Info 배송정보

출고 정보 ⓘ 결제 3일 이내 출고 ▾
배송 방법 국내 배송 / 무신사 직배송 / CJ대한통운

Price Info 가격정보

무신사 판매가 74,900원
무신사 회원가 71,904 ~ 74,900원 ▾
무이자 할부가 8,988 ~ 9,363원 (8개월)
무신사 적립금 최대 2,996원
시즌 ⓘ / 성별 2020 F/W 남
조회수(1개월) 1.7만 회 이상
누적판매(1년) 5.9천 개 이상 (결제완료-반품)
좋아요 10,958
구매 후기 ⓘ ★★★★★ 4.8 / 후기 1,969개 보기
#오버핏셔츠 #루즈핏셔츠 #가을셔츠 #남자셔츠 #울체크셔츠
#오버핏남방 #글렌체크셔츠

19~28세, 남성 에게 인기 많은 상품 ▾

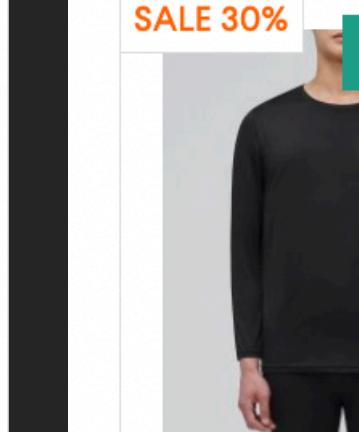
Delivery Info 배송정보

출고 정보 ⓘ 예약배송 05.02 출고
배송 방법 국내 배송 / 일정사 배송

Price Info 가격정보

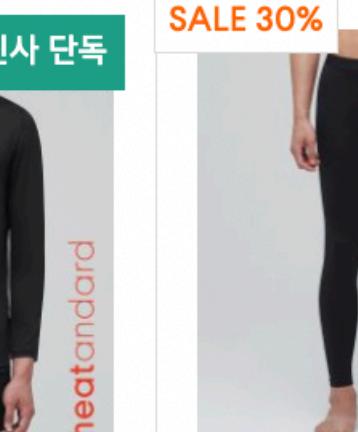
무신사 판매가 48,000원
무신사 회원가 35,136 ~ 38,400원 ▾ 20% 이상 할인
무신사 적립금 최대 1,636원

무신사스토어는 전 상품 무료배송입니다. 회원 특별 혜택



SALE 30%
무신사 단독
heatandard

무신사 스탠다드
핫탠다드 맨즈 크루 넥 티셔츠 [블랙]
13,900원 9,690원
MEMBERSHIP PRICE 5,084
2,385



SALE 30%
무신사 단독
heatandard

무신사 스탠다드
핫탠다드 맨즈 타이즈 [블랙]
13,900원 9,690원
MEMBERSHIP PRICE 7,038
3,351



SALE 30%
무신사 단독
heatandard

무신사 스탠다드
핫탠다드 우먼즈 크루 넥 티셔츠 [블랙]
13,900원 9,690원
MEMBERSHIP PRICE 2,003
2,344



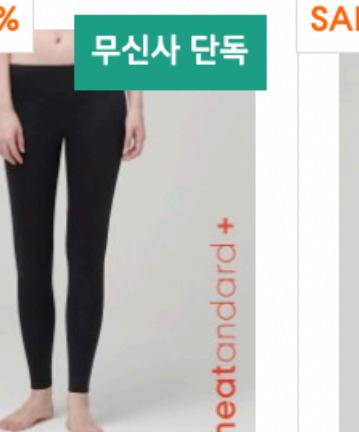
SALE 30%
무신사 단독
heatandard+

무신사 스탠다드
핫탠다드 플러스 우먼즈 크루 넥 티셔츠 [블랙]
17,900원 12,490원
MEMBERSHIP PRICE 268
443



SALE 30%
무신사 단독
heatandard

무신사 스탠다드
핫탠다드 플러스 우먼즈 타이즈 [블랙]
17,900원 12,490원
MEMBERSHIP PRICE 234
346



SALE 30%
무신사 단독
heatandard

무신사 스탠다드
핫탠다드 우먼즈 크루 넥 티셔츠 [다크 그레이]
13,900원 9,690원
MEMBERSHIP PRICE 214
136

비지 > 슬acks / 허드 (무신사 스텠다드)

무신사 단독 글렌 체크 테이퍼드 히든 밴딩 크롭 슬랙스 [미디엄 그레이] GLEN CHECK TAPERED HIDDEN BANDING CROP SLACKS[MEDIUM GREY]



무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

브랜드 / 품번 MUSINSA STANDARD / MMJPS507-MR
시즌 ⓘ / 성별 2020 F/W 남
조회수(1개월) 1.4천 회 이상
누적판매(1년) 400개 이상 (결제완료-반품)
좋아요 1,415
구매 후기 ⓘ ★★★★★ 4.9 / 후기 679개 보기
#트리우저 #정장바지 #수트팬츠 #체크팬츠 #체크
#피페트슬랙스 #무던파업

24~33세, 남성 에게 인기 많은 상품 ▾

Delivery Info 배송정보

출고 정보 ⓘ 결제 3일 이내 출고 ▾
배송 방법 국내 배송 / 무신사 직배송 / CJ대한통운

Price Info 가격정보

무신사 판매가 34,900원
무신사 회원가 33,504 ~ 34,900원 ▾
무신사 적립금 최대 1,396원

2022년도 1분기 리뉴얼 추천상품

ITEM RECOMMENDATION

유사도 바탕 최종 상품 추천을 위한 아이템 추천

상의 > 니트/스웨터 (무신사 스탠다드)
무신사 단독 [클랜디드] 미니멀 니티드 폴로 셔츠 [블랙] | MINIMAL KNITTED POLO SHIRT [BLACK]

무신사 단독

무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD / MMAWS306-BK
- 시즌 ⓘ 성별 2021 S/S 남
- 조회수(개월) 700회 이상
- 누적판매(1년) 300개 이상 (결제완료-반품)
- 좋아요 346
- 구매 후기 ⓘ ★★★★★ 4.9 후기 295개 보기
- #반팔니트 #어辱니트 #폴로니트 #칼라니트 #니트컬렉션
#폴로셔츠컬렉션 #클랜디드 #무덴피업

24-33세, 남성에게 인기 많은 상품 ▶

Delivery Info 배송정보

- 출고 정보 ⓘ 결제 3일 이내 출고 ▶
- 배송 방법 국내 배송 / 무신사 직배송 CJ대한통운

Price Info 가격정보

- 무신사 판매가 53,900원
- 무신사 회원가 51,744 ~ 50,000원

상의 > 피케/카라 티셔츠 (무신사 스탠다드)
무신사 단독 슬릿 넥 니티드 폴로 셔츠 [블랙] | SLIT NECK KNITTED POLO SHIRT [BLACK]

무신사 단독

무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD / MMJWP202-BK
- 시즌 ⓘ 성별 2020 S/S 남
- 조회수(개월) 1,989회 이상
- 누적판매(1년) 500개 이상 (결제완료-반품)
- 좋아요 980
- 구매 후기 ⓘ ★★★★★ 4.7 후기 955개 보기
- #반팔니트 #반팔셔츠 #카라티 #클랜디드 #폴로셔츠컬렉션
#피케셔츠 #니트컬렉션

24-33세, 남성에게 인기 많은 상품 ▶

Delivery Info 배송정보

- 출고 정보 ⓘ 결제 3일 이내 출고 ▶
- 배송 방법 국내 배송 / 무신사 직배송 CJ대한통운

Price Info 가격정보

- 무신사 판매가 53,900원

모자 > 헤/야구 모자 (무신사 스탠다드)
무신사 단독 라이트웨이트 폴딩 캠 [에시드 그린] | LIGHTWEIGHT FOLDING CAP [ACID GREEN]

무신사 단독

무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD / MICP0003-AE
- 시즌 ⓘ 성별 2020 ALL 남
- 조회수(개월) 400회 이상
- 좋아요 32
- 구매 후기 ⓘ ★★★★★ 4.4 후기 442개 보기
- #기능성 #라이트웨이트컬렉션 #방수 #아웃도어 #캠프캡
#라이트웨이트 #무덴피업

Delivery Info 배송정보

- 출고 정보 ⓘ 결제 3일 이내 출고 ▶
- 배송 방법 국내 배송 / 무신사 직배송 CJ대한통운

Price Info 가격정보

- 무신사 판매가 15,900원
- 무신사 회원가 15,264 ~ 15,900원
- 무신사 적립금 최대 636원

무신사스토어는 전 상품 무료배송입니다. 회원 특별 혜택

최저가 경쟁기록 달리기 설정해 보시고 알리세요!

상의 > 니트/스웨터 (무신사 스탠다드)
무신사 단독 [클랜디드] 미니멀 쇼트 슬리브 크루넥 니트 [블랙] | MINIMAL SHORT SLEEVE CREW NECK KNIT [BLACK]

무신사 단독

무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD / MMAWS304-BK
- 시즌 ⓘ 성별 2021 S/S 남
- 조회수(개월) 800회 이상
- 누적판매(1년) 200개 이상 (결제완료-반품)
- 좋아요 309
- 구매 후기 ⓘ ★★★★★ 4.8 후기 204개 보기
- #반팔니트 #라운드넥니트 #여름니트 #유넥 #유네니트
#니트컬렉션 #클랜디드

24-33세, 남성에게 인기 많은 상품 ▶

Delivery Info 배송정보

- 출고 정보 ⓘ 결제 3일 이내 출고 ▶
- 배송 방법 국내 배송 / 무신사 직배송 CJ대한통운

Price Info 가격정보

- 무신사 판매가 49,900원
- 무신사 회원가 47,904 ~ 49,900원
- 무신사 적립금 최대 1,996원

상의 > 니트/스웨터 (커스텀어글리)
한정 판매 줄 소프트 숏 슬리브 크루넥 니트 블랙 | COOL SOFT SHORT SLEEVE CREW NECK KNIT BLACK

한정 판매

무신사 스토어 포함한 제한된 판매처에서만 구매 가능한 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 COSTUME O'CLOCK / COOSKTO10BLACK
- 시즌 ⓘ 성별 2020 S/S 남
- 조회수(개월) 3,191회 이상
- 누적판매(1년) 400개 이상 (결제완료-반품)
- 좋아요 408
- 구매 후기 ⓘ ★★★★★ 4.8 후기 1,652개 보기
- #기본니트 #라운드니트 #반팔니트 #베이직니트 #스탠다드
#여름니트

24-33세, 남성에게 인기 많은 상품 ▶

Delivery Info 배송정보

- 출고 정보 ⓘ 결제 3일 이내 출고 ▶
- 배송 방법 국내 배송 / 입점자 배송 CJ대한통운

Price Info 가격정보

- 무신사 판매가 45,000원
- 무신사 회원가 28,822 ~ 31,500원
- 무신사 적립금 최대 1,260원

2022년도 2분기 리뉴얼 추천상품

ITEM RECOMMENDATION

유사도 바탕 최종 상품 추천을 위한 아이템 추천

[마우디 > 블루종/MA-1 \[무신사 스텠다드\]](#)

무신사 단독 [유아인 착용 상품] 크링클 오버사이즈 MA-1 보머 재킷 [라이트 그레이] | CRINKLE OVERTSIZE MA-1 BOMBER JACKET [LIGHT GREY]

무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD / MMBUJ104-LG
- 시즌 ② / 성별 2022 S/S 남
- 조회수(개월) 1.4만 회 이상
- 누적판매(1년) 100개 이상 (결제완료-반품)
- 좋아요 673
- 구매 후기 ② ★★★★★ 4.9 후기 61개 보기
- #점퍼 #항공점퍼 #밀리터리 #방수재킷 #캐주얼아우터
#한정판 #마운틴 #오버핏 #유아인착상품

24~33세, 남성에게 인기 많은 상품 ▾

Delivery Info 배송정보

- 출고 정보 ② 결제 3일 이내 출고 ▾
- 배송 방법 국내 배송 무신사 직배송 CJ대한통운

Price Info 가격정보

- 무신사 판매가 85,900원

[무신사 단독 > 슈트 팬츠/슬랙스 \(무신사 스텠다드\)](#)

무신사 단독 라이트웨이트 와이드 하든 밴딩 슬랙스 [라이트 그레이] | LIGHTWEIGHT WIDE HIDDEN BANDING SLACKS [LIGHT GREY]

무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD MMJPS204-LG
- 시즌 ② / 성별 2020 S/S 남
- 조회수(개월) 3.3천 회 이상
- 누적판매(1년) 400개 이상 (결제완료-반품)
- 좋아요 862
- 구매 후기 ② ★★★★★ 4.8 후기 8,763개 보기
- #라이트웨이트컬렉션 #셋업 #퍼펙트슬랙스 #하리밴딩 #정장바지
#여름슬랙스 #와이드핏 #라이트웨이트

19~28세, 남성에게 인기 많은 상품 ▾

Delivery Info 배송정보

- 출고 정보 ② 결제 3일 이내 출고 ▾
- 배송 방법 국내 배송 무신사 직배송 CJ대한통운

Price Info 가격정보

- 무신사 판매가 32,900원
- 무신사 회원가 31,584 ~ 32,900원 ▾
- 무신사 적립금 최대 1,316원

[SALE 40% 무신사 단독](#)

[SALE 40% 무신사 단독](#)

[SALE 50% 무신사 단독](#)

무신사 스텠다드 코튼 오버사이즈 맥 코트 [크림]
무신사 스텠다드 코튼 오버사이즈 맥 코트 [블랙]
무신사 스텠다드 코튼 오버사이즈 맥 코트 [크림]

89,900원 53,890원
89,900원 53,890원
89,900원 44,890원

MEMBERSHIP PRICE 603
MEMBERSHIP PRICE 849
MEMBERSHIP PRICE 47

1,541
2,590
255

[스니커즈 > 캐비스/단화 \(무신사 스텠다드\)](#)

무신사 단독 독일군 스니커즈 [오프 화이트] | GERMAN ARMY SNEAKERS [OFF WHITE]

Ortholite® comfort foam insoles

무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD MMJFS007-OW
- 시즌 ② / 성별 2020 ALL 남, 여
- 조회수(개월) 1.5만 회 이상
- 누적판매(1년) 2,617 개 이상 (결제완료-반품)
- 좋아요 5,047
- 구매 후기 ② ★★★★★ 4.8 후기 1,339개 보기
- #트레이너 #레플리카 #커플신발 #오슬라이트 #사계절 #운동화
#유니섹스 #무던직업 #레더

19~28세, 남성에게 인기 많은 상품 ▾

Delivery Info 배송정보

- 출고 정보 ② 결제 3일 이내 출고 ▾
- 배송 방법 국내 배송 무신사 직배송 CJ대한통운

Price Info 가격정보

- 무신사 판매가 69,900원
- 무신사 회원가 67,104 ~ 69,900원 ▾
- 무이자 할부가 8,388 ~ 8,738원 (3개월)
- 무신사 적립금 최대 2,796원

[무신사 단독](#)

[무신사 단독](#)

[무신사 단독](#)

무신사 단독 티셔츠 [무신사 스텠다드]

무신사 단독 마초의 사춘기 반팔 티셔츠 [블랙] | MACHO'S SACHUNKI T-SHIRT [BLACK]

무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD CWBTS105-BK
- 시즌 ② / 성별 2022 S/S 남, 여
- 조회수(개월) 1.3천 회 이상
- 누적판매(1년) 50개 이상 (결제완료-반품)
- 좋아요 91
- 구매 후기 ② ★★★★★ 4.8 후기 4개 보기
- #면서초 #무신사로드트립 #그래픽티 #라운드넥 #데글러프
#MU-T

34세 이상, 남성에게 인기 많은 상품 ▾

Delivery Info 배송정보

- 출고 정보 ② 결제 3일 이내 출고 ▾
- 배송 방법 국내 배송 무신사 직배송 CJ대한통운

Price Info 가격정보

- 무신사 판매가 21,900원
- 무신사 회원가 21,024 ~ 21,900원 ▾
- 무신사 적립금 최대 876원

2022년도 3분기 리뉴얼 추천상품

ITEM RECOMMENDATION

유사도 바탕 최종 상품 추천을 위한 아이템 추천

아우터 > 겨울 기타 코트 (무신사 스탠다드)

무신사 단독 캐시미어 블렌드 더플 코트 [브라운] | CASHMERE BLEND DUFFLE COAT [BROWN]

무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD / MMJCT706-BR
- 시즌 Ⓛ / 성별 2021 F/W / 남
- 조회수(1개월) 2.4천 회 이상
- 누적판매(1년) 50개 이상 (결제완료-반품)
- 좋아요 407
- 구매 후기 Ⓛ ★★★★☆ 4.9 후기 636개 보기
- #롱코트 #울코트 #딱볶이코트 #루즈핏 #기본코트 #코트컬렉션

19~28세, 남성에게 인기 많은 상품 ▶

Delivery Info 배송정보

- 출고 정보 Ⓛ 결제 3일 이내 출고 ▶
- 배송 방법 국내 배송 / 무신사 직배송 / CJ대한통운

Price Info 가격정보

무신사 판매가 179,900원

무신사 회원가 112,214 ~ 116,890원 ▶ 35% 이상 할인

무이자 할부가 14,027 ~ 14,611원 (8개월)

무신사 적립금 최대 4,676원

아우터 > 겨울 기타 코트 (무신사 스탠다드)

무신사 단독 더플 코트 [카멜] | DUFFLE COAT [CAMEL]

무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD / MMJCT716-CA
- 성별 남
- 조회수(1개월) 3.7천 회 이상
- 누적판매(1년) 100개 이상 (결제완료-반품)
- 좋아요 791
- 구매 후기 Ⓛ ★★★★☆ 4.8 후기 117개 보기
- #남자코트 #롱코트 #울코트 #딱볶이코트 #루즈핏 #코트컬렉션

19~28세, 남성에게 인기 많은 상품 ▶

Delivery Info 배송정보

- 출고 정보 Ⓛ 결제 3일 이내 출고 ▶
- 배송 방법 국내 배송 / 무신사 직배송 / CJ대한통운

Price Info 가격정보

무신사 판매가 179,900원

무신사 회원가 112,214 ~ 116,890원 ▶ 35% 이상 할인

무이자 할부가 14,027 ~ 14,611원 (8개월)

무신사 적립금 최대 4,676원

아우터 > 슈트/블레이저 재킷 (무신사 스탠다드)

무신사 단독 [유아인 착용 상품] 빅 대디 오버사이즈 블레이저 [블랙] | BIG DADDY OVERTSIZE BLAZER [BLACK]

무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD / MMAJK009-BK
- 시즌 Ⓛ / 성별 2022 S/S 남
- 조회수(1개월) 6.2만 회 이상
- 누적판매(1년) 1.4천 개 이상 (결제완료-반품)
- 좋아요 2,409
- 구매 후기 Ⓛ ★★★★☆ 4.9 후기 367개 보기
- #오버핏 #셋업 #셔링 #파렛츠블레이저 #블레이저 #한정판아우터 #유아인착용상품

19~28세, 남성에게 인기 많은 상품 ▶

Delivery Info 배송정보

- 출고 정보 Ⓛ 결제 3일 이내 출고 ▶
- 배송 방법 국내 배송 / 무신사 직배송 / CJ대한통운

Price Info 가격정보

무신사 판매가 83,900원

무신사 회원가 80,544 ~ 83,900원 ▶

무이자 할부가 10,068 ~ 10,488원 (8개월)

무신사 적립금 최대 3,356원

보디 > 포인트 메이크업 > 립밤 (무신사 스탠다드)

무신사 단독 무신사 스탠다드 시어 버터 립밤 3.8g 3팩 | MUSINSA STANDARD SHEA BUTTER LIP BALM 3.8g 3PACK

무신사 단독
무신사 스토어에서만 단독 판매하는 상품입니다.

Product Info 제품정보

- 브랜드 / 품번 MUSINSA STANDARD / MMJOC009-BK
- 성별 남, 여
- 조회수(1개월) 1000회 이상
- 누적판매(1년) 600개 이상 (결제완료-반품)
- 좋아요 706
- 구매 후기 Ⓛ ★★★★☆ 4.8 후기 466개 보기
- #립케어 #무탄피압 #립밤 #남자립밤 #보습립밤 #남자립밤추천 #시아바터립밤 #크스메틱컬렉션 #홀리데이 #세트상품

19~28세, 남성에게 인기 많은 상품 ▶

Delivery Info 배송정보

- 출고 정보 Ⓛ 결제 3일 이내 출고 ▶
- 배송 방법 국내 배송 / 무신사 직배송 / CJ대한통운

Price Info 가격정보

무신사 판매가 13,900원

무신사 회원가 11,990 ~ 12,490원 ▶ 10% 이상 할인

무신사 적립금 최대 500원

무신사 단독

무신사 스탠다드

스트라이프 보트 넥 긴팔 티셔츠 [아이보리/블랙] 19,900원

MEMBERSHIP PRICE ▾

★★★★★ 1,439
2,672

무신사 단독

무신사 스탠다드

스트라이프 보트 넥 긴팔 티셔츠 [아이보리/네이비] 19,900원

MEMBERSHIP PRICE ▾

★★★★★ 457
1,284

2022년도 4분기 리뉴얼 추천상품

USING TECHNIQUE DESCRIPTION

TF - IDF

[TF-IDF 포함] 전체 과정 함수화

```

def get_title(num):
    url = 'https://www.musinsa.com/app/goods/' + str(num)

def vect_count(data1, num):
    tfidf = TfidfVectorizer(tokenizer = lambda x: text_cleaning(x, stopwords[num]), max_features=2000, min_df=4, max_df=0.1)
    tf_idf_vect = tfidf.fit_transform(data1['상품설명'])
    tf_idf_vect

hangul = re.compile('[^ㄱ-ㅎ | 가-힣]') # 정규 표현식 처리
result = hangul.sub('', text)
okt = Okt() # 형태소 추출
nouns = okt.nouns(result)
nouns = [x for x in nouns if len(x) > 1] # 한글자 키워드 제거

if len(text) > 10000:
    nouns = [x for x in nouns if x not in stopwords]
else:
    nouns = [x for x in nouns if x not in product_stopwords]
return nouns

def vect_count(data1, num):
    tfidf = TfidfVectorizer(tokenizer = lambda x: text_cleaning(x, stopwords[num]), max_features=2000, min_df=4, max_df=0.1)
    tf_idf_vect = tfidf.fit_transform(data1['상품설명'])
    tf_idf_vect

similarity_simple_pair = cosine_similarity(tf_idf_vect[0], tf_idf_vect)

sorted_index = similarity_simple_pair.argsort()[:, ::-1] # 인덱스 내림차순으로 저장
sorted_index = sorted_index[:, 1:]
print(sorted_index.reshape(-1)[:10])

curating_sim_value = np.sort(similarity_simple_pair.reshape(-1))[:, ::-1] # [:,-1]은 처음부터 끝까지 역순으로 정렬
curating_sim_value = curating_sim_value[1:]

curating_sim_df = pd.DataFrame()
curating_sim_df['상품코드'] = data1.iloc[sorted_index.reshape(-1)[10].tolist()]['상품번호']
curating_sim_df['유사도'] = curating_sim_value[:10]
print(curating_sim_value[:10])

curating_sim_df['상품코드'] = curating_sim_df['상품코드'].astype(str)
sns.barplot(x = '유사도', y = '상품코드', data = curating_sim_df);
plt.show()

return curating_sim_df['상품코드']

```

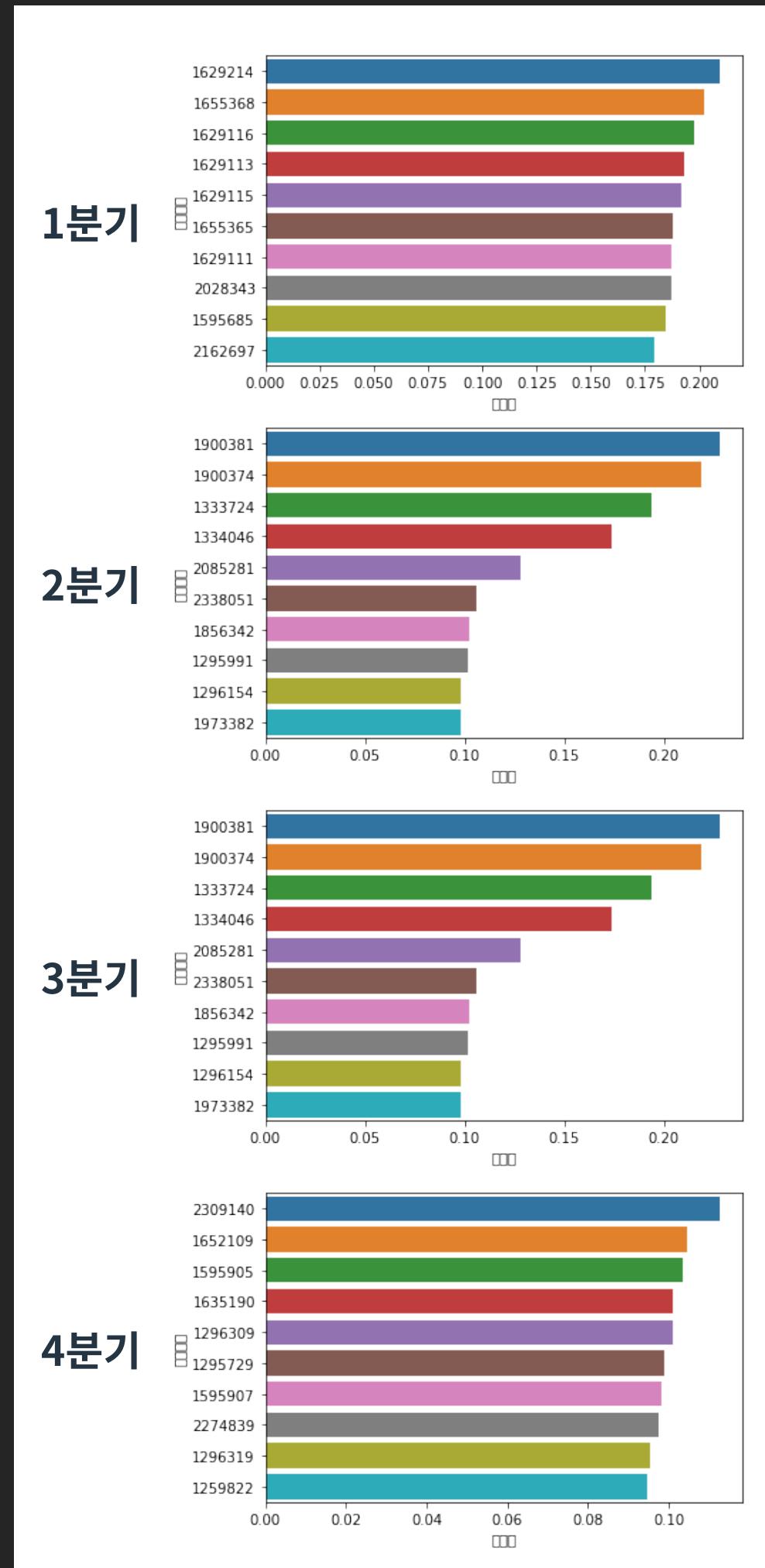
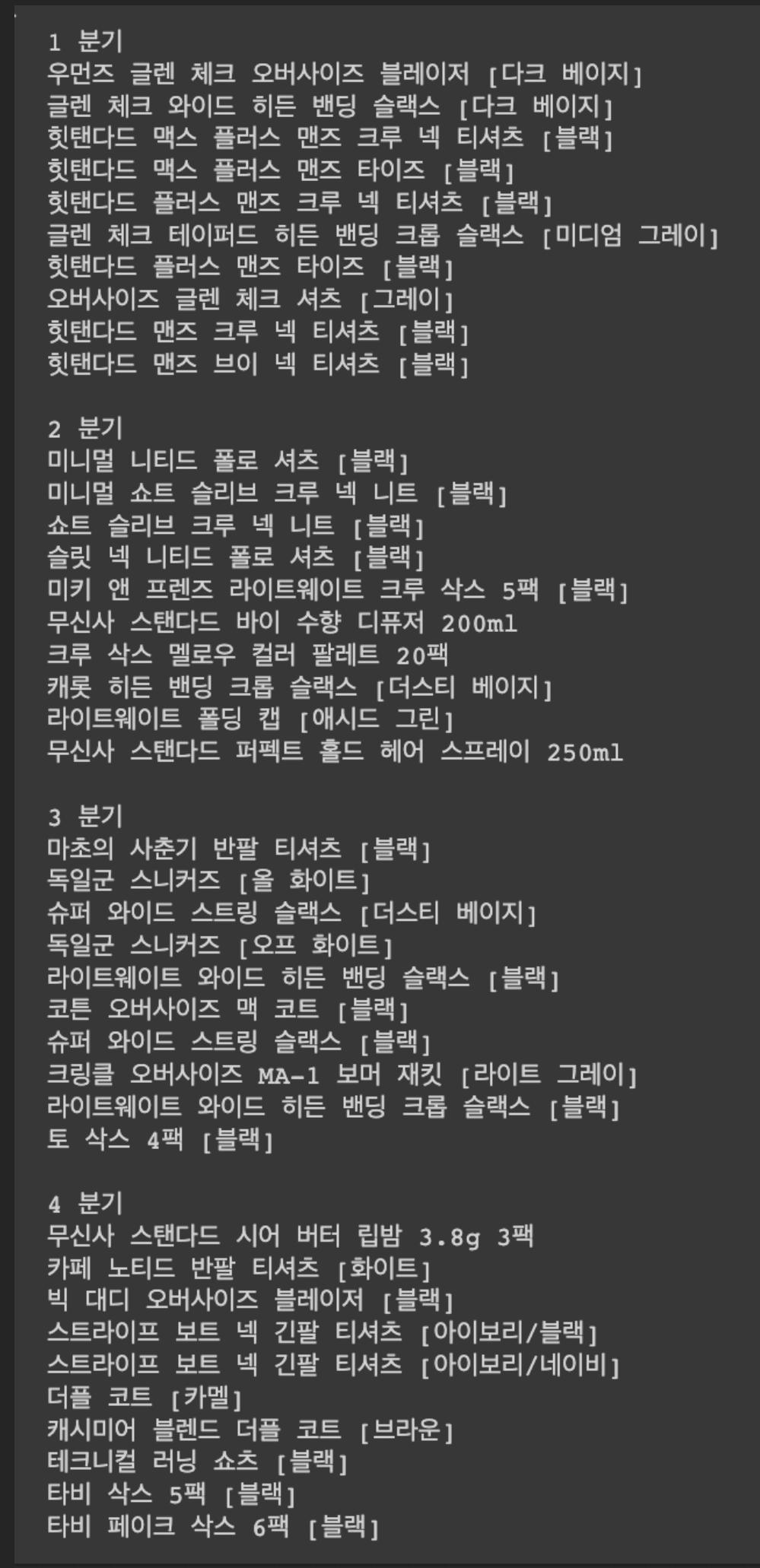
3가지 TF-IDF 파라미터 값을 변경해가며,

가장 올바르게 분석된 값을 찾아보았다.

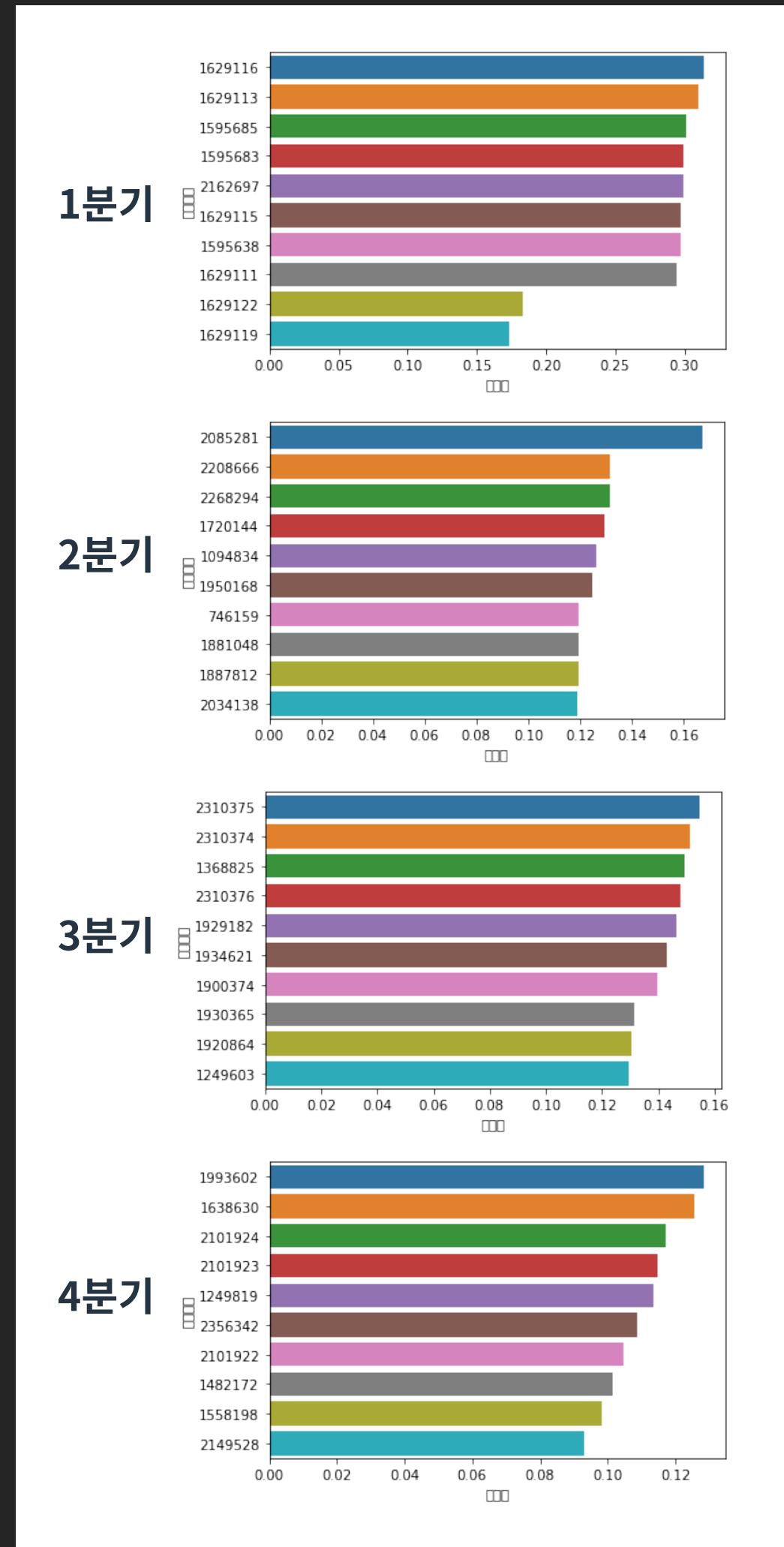
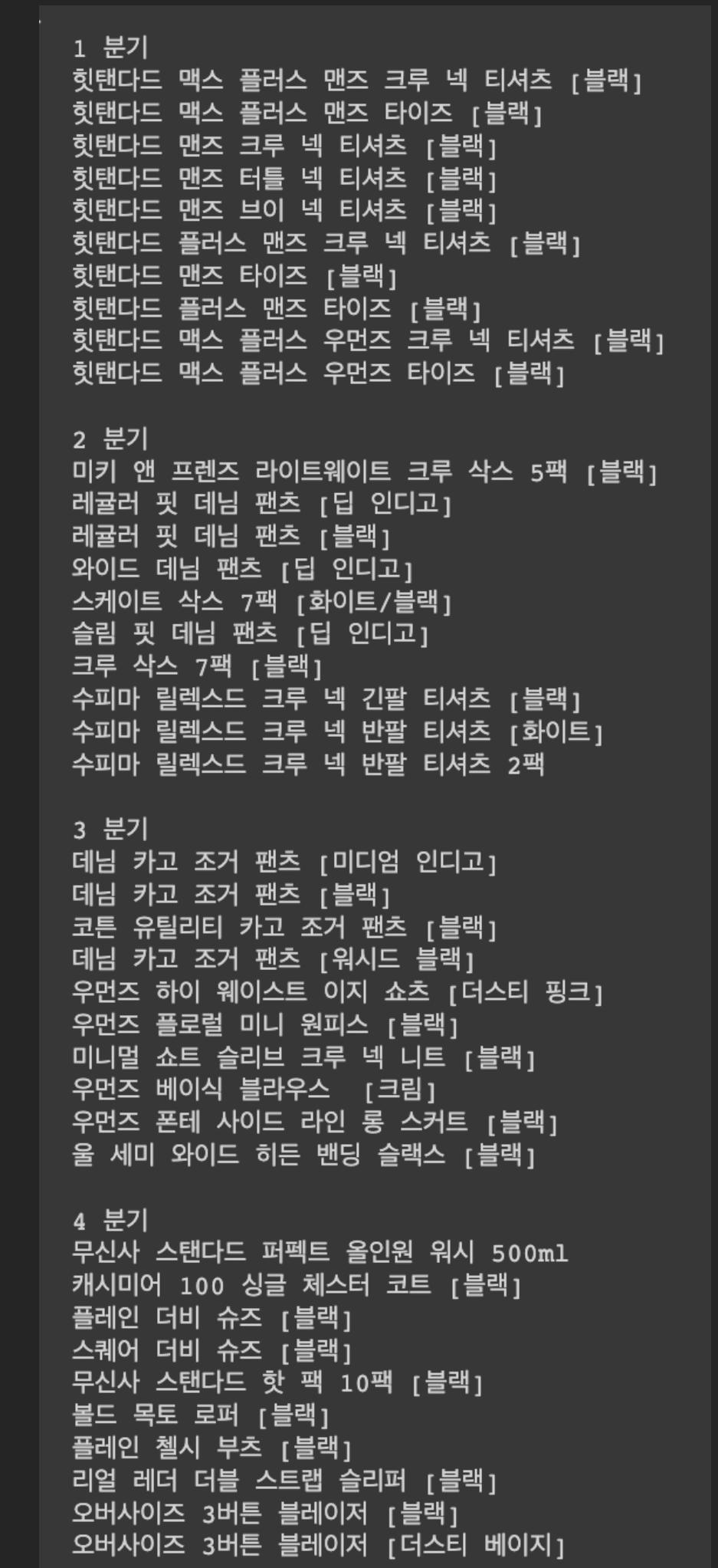
- max_features
- min_df
- max_df

USING TECHNIQUE DESCRIPTION

max_features=1500, min_df=3, max_df=0.1



max_features=1000, min_df=10, max_df=0.3



PREDICTING REVIEW STAR RATE

MUSINSA 입점브랜드에 관한 실질적 평가지표 제시

ISSUE POINT

부정 댓글 데이터셋의 부재

2022년도 6월까지의 댓글

+

2021년도 1,2,3점 (부정 댓글) 추가 크롤링

Star score "1":	463
Star score "2":	563
Star score "3":	5332
Star score "4":	39242
Star score "5":	298540

‘무신사 > 회원후기’ 데이터셋

TEXT AUGMENTATION

KoEDA [Easy Data Augmentation for Korean]

단어를 삽입 / 삭제 / 위치 변경 / 유의어로 대체 하여 증식하는 기법

Back Translation

기존 텍스트를 외국어로 번역한 뒤, 다시 기존의 언어로 번역하여
증식하는 기법

Generation

키워드의 앞, 뒤 상관관계 및 유사도를 기반하여 글자를 생성하여
증식하는 방법



< STAR SCORE 1,2 >
Text Augmentation

CHALLENGE WAY

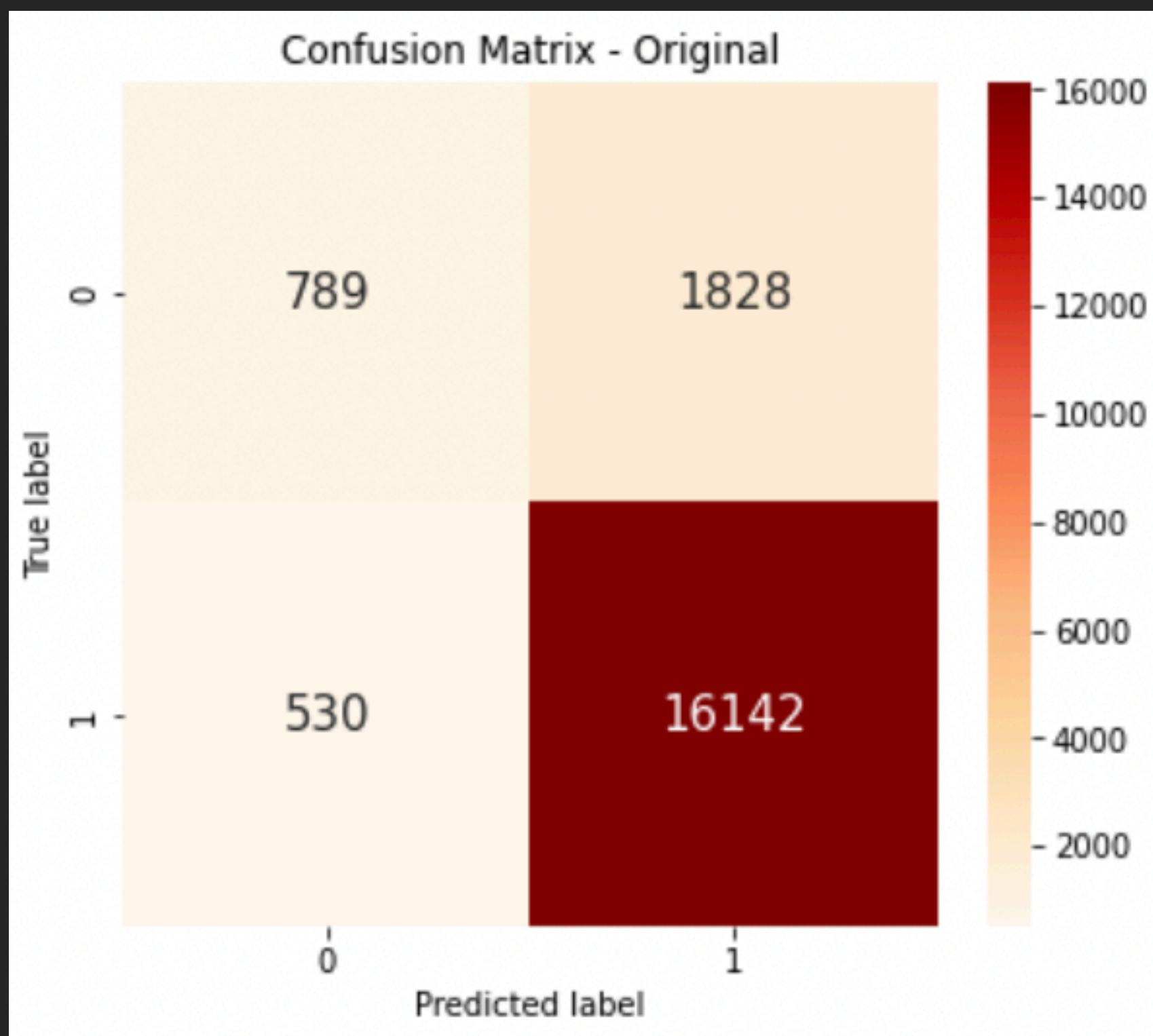
라벨 수작업

SCORE_1000 단어 점수 세분화

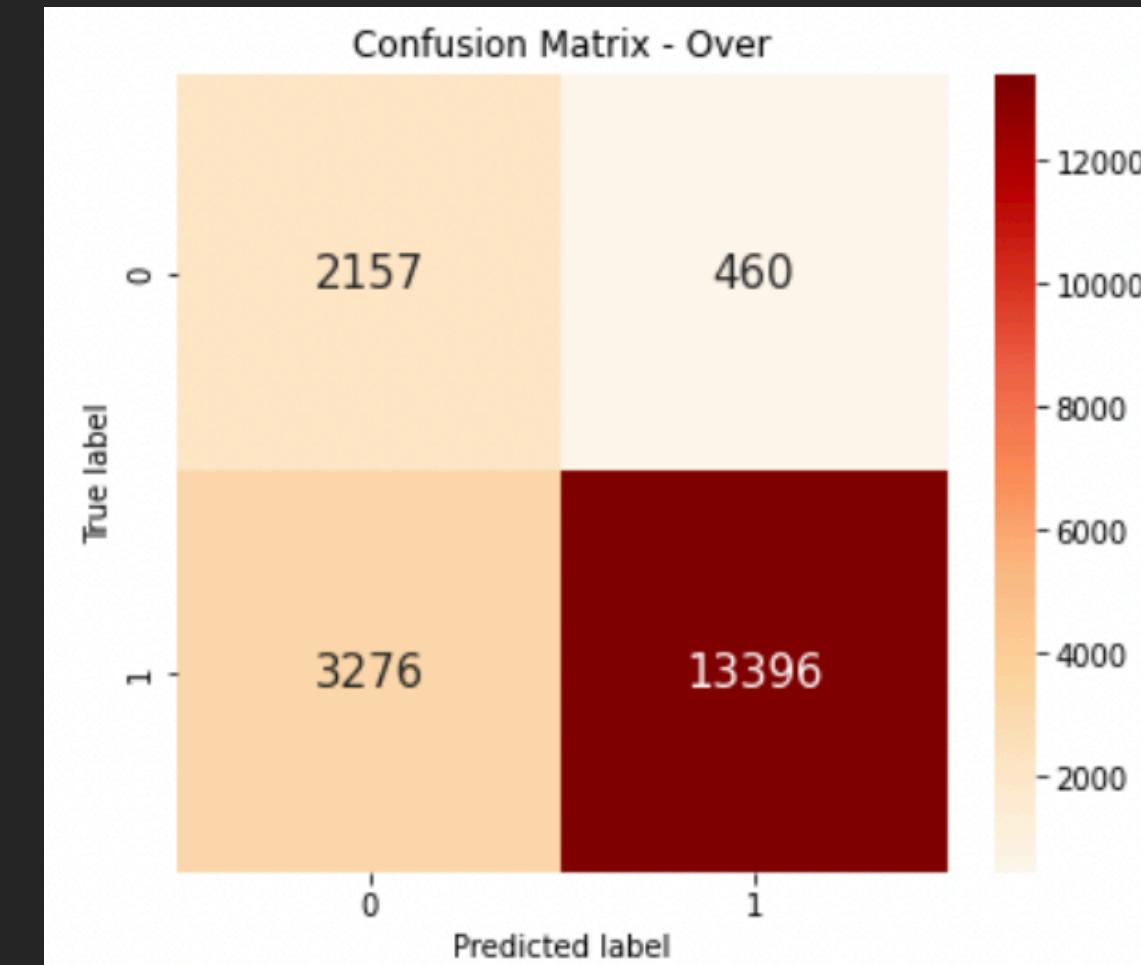
	긍정단어	실제 리뷰 평가(긍정)	부정단어	실제 리뷰 평가(부정)
0	압도	3	부각	-2
1	다행	1	자꾸	-1
2	완전	3	환불	-3
3	다른	1	살바	-2
4	정말	3	한철	-1
5	매우	3	오염	-3
6	성공	3	심해	-3
7	번창	2	실망	-3
8	핏도좋	2	구김	-2
9	평상복	1	짜증	-3
10	이쁜핏	3	싼마이	-2
11	힙하	2	흐물거리	-1
12	감동	3	의문	-1
13	힙해	2	별로	-2
14	약속	1	자모	0

CHALLENGE WAY

성능 개선을 위한 SAMPLING 작업



Original

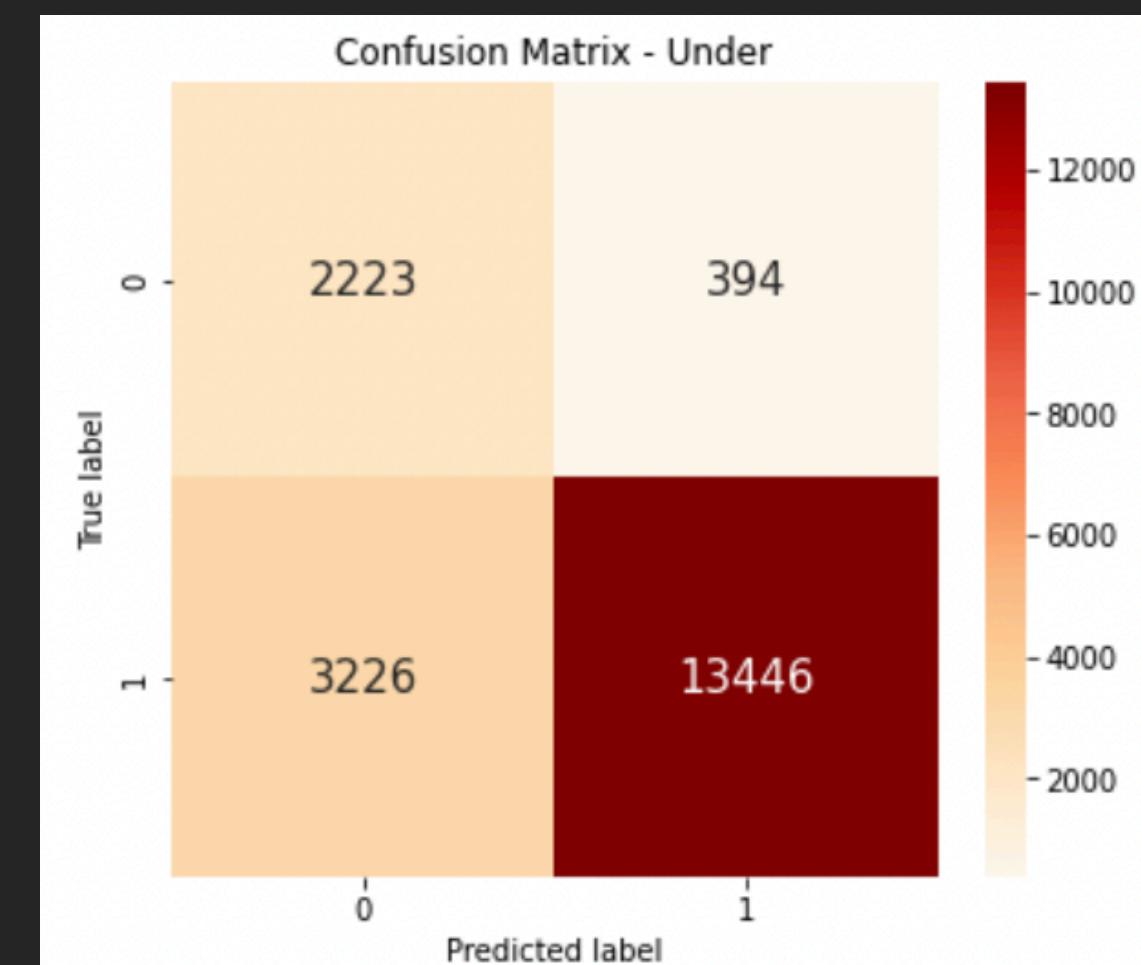


OVER - SAMPLING

SMOTE 를 사용하여 긍정 댓글 수에 맞추어 Sampling

SMOTE 란?

[Synthetic Minority Over-sampling Technique]
낮은 비율로 존재하는 클래스의 데이터를
최근접 이웃[k-NN 알고리즘] 알고리즘을 활용하여
새롭게 생성하는 방법



UNDER - SAMPLING

Text Augmentation 을 하지 않은 상태에서
부정 댓글 수에 맞추어 Sampling

TARGET

Let's think about it in terms of MUSINSA



REVIEW

실질적인 판매량에 영향력

좋은 평점 리뷰

쇼핑몰의 이익에 영향

사용 후기가 시장을 움직인다... 전문가 입 김보다 쎈 '유저스 코멘트'

<https://www.asiatoday.co.kr/view.php?key=20200219010011433>

커머스 업계에서 고객 후기의 중요성이 커지고 있다. 아마존, 네이버 쇼핑, 쿠팡, 11번가와 같은 대형 오픈마켓 커머스들 또한 리뷰어 랭킹제 도입, 리뷰 항목 체계화, 파격적인 리워드 제공 등을 통해 나름의 리뷰 시스템을 활발하게 구축하고 있다. 이에 발 맞춰 온라인 쇼핑몰이 사용하는 리뷰 솔루션 업체들 역시 서비스를 빠르게 고도화 중이다.

이처럼 커머스들이 너나 할 것 없이 리뷰 시스템에 투자하는 이유는 소비자 후기가 매출에 실질적인 영향을 주기 때문이다. '맥킨지앤컴퍼니'의 조사에 따르면 실구매자의 후기는, 쇼핑몰의 추천보다 잠재소비자의 구매 결정에 10배 가량 높은 영향력을 행사한다고 한다.

<https://www.joongang.co.kr/article/21404190#home>

리뷰와 판매 매출 관계

제품 리뷰는 온라인 쇼핑 판매 매출에 직접적인 영향을 미칩니다. 우선 수많은 판매자가 있는 온라인 시장에서 해당 판매자의 신뢰가 올라갑니다. 판매 사기도 많은 만큼 같은 제품이라도 리뷰가 많은 판매자는 더 안전하다고 생각이 듭니다.

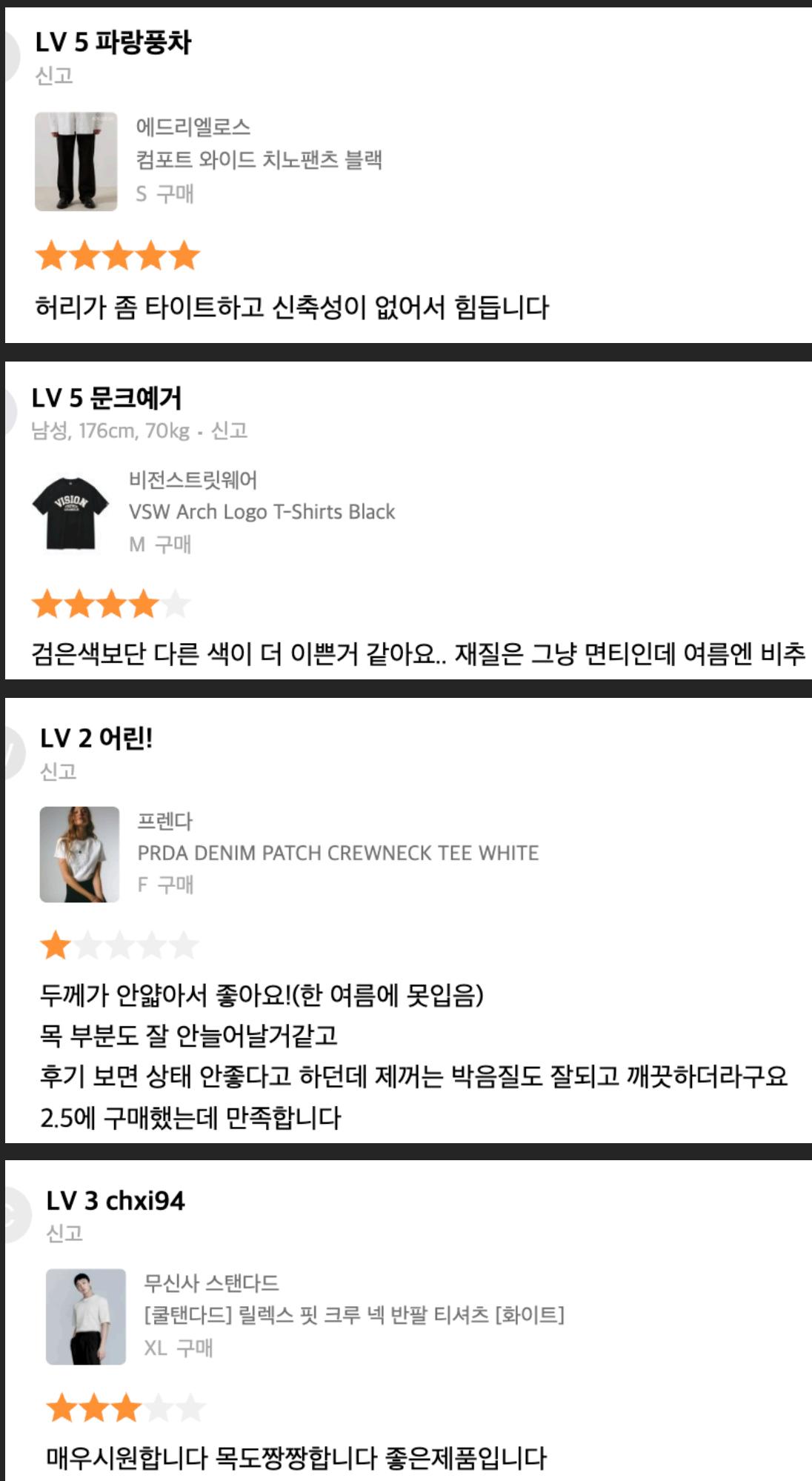
리뷰는 제품 자체에 대한 믿음도 올려줍니다. 일단 내용의 좋고 나쁨을 떠나 리뷰가 많다는 것은 많은 소비자가 구매를 했다는 것이므로 구매조차 하지 않는 제품보다는 믿음이 가는 것은 확실합니다. 리뷰가 일정 수준 이상 있고 다음이 전체 사용 리뷰의 평이 좋냐 나쁘냐가 고려 요인이 될 것입니다.

일정 수준 리뷰가 쌓여 있다면 좋은 리뷰가 많을수록 이를 본 소비자가 그 제품을 구매할 가능성이 커지게 됩니다. 일종의 동료 압박이나 반복 학습의 영향이라고 말할 수 있습니다. 단순히 많은 사람이 좋다고 하는 제품은 실제 써보지 않아도 좋아 보이는 것입니다.

<https://revlab.tistory.com/entry/%EC%A0%9C%ED%92%88-%EB%A6%AC%EB%B7%BO-%ED%9A%A8%EA%B3%BC-%ED%95%9C%EA%B3%84>

TARGET

보다 정확한 별점 점수를 예측하기 위한 모델 제작



내용은 좋지 않지만
별점은 높은 경우

내용은 좋은데
별점은 낮은 경우



MUSINSA 입점브랜드에 관한 실질적 평가지표를 제시한다.

저평가된 브랜드와 고평가 된 브랜드를 구분

고객 만족도 기반 해당 브랜드와의 계약 연장성에 대한 고찰

WORKFLOW



DATASET

MUSINSA

STAR RATE 5
STAR RATE 4
STAR RATE 3
STAR RATE 2
STAR RATE 1

+

NAVER

STAR RATE 2
STAR RATE 1

<부정 댓글 데이터셋 부재>
비단 무신사에만 있는 데이터가 아니다.
네이버 쇼핑에서 데이터셋 추가적으로 수집

AUGMENTATION

STAR RATE 3
STAR RATE 2
STAR RATE 1

MODEL [SEARCH]

BERT (Multilingual) : GOOGLE 개발

> 목적

- 다양한 언어에 폭 넓게 적용
- Character-level (글자, 알파벳) 단위를 사용한 설계



한계점

- 과도하게 큰 model의 크기
- non-English downstream task에서의 성능부진

> 배경 및 특징

- 104개의 language의 Wikipedia dataset으로 학습된 model
- 알파벳을 사용하는 언어들은 전체 character가 26개에 불과
- 데이터가 풍부한 언어 [ex. German / French] 의 경우,
Wikipedia + Add Dataset로 pretrain을 수행
- 자연언어를 양방향으로 사전학습하는 모델

MODEL [SEARCH]

ABOUT KOREAN

“ 한국어 (KOREAN) ”

< 교착어 >

어근과 접사에 의해 단어의 기능이 결정되는 언어의 형태

> 형태소

English / French / German (굴절어) < **한국어 (교착어)**

> 특징

- 음절 기반 = 약 11,172개 character가 존재
- BERT에서는 이 중 오직 1,187개의 character만이 포함

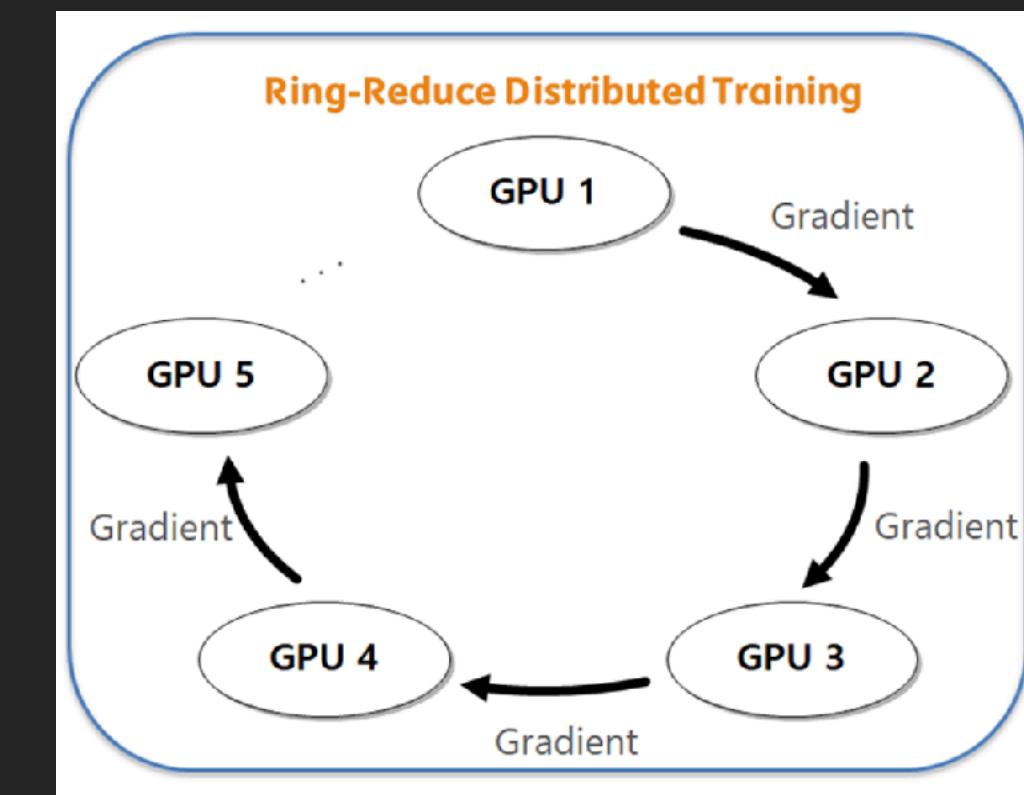
MODEL [CHOICE]

KoBERT : T - BRAIN (SKT) 개발

“ KOREAN + BERT ”

> 한국어의 불규칙한 언어 변화의 특성을 반영

- 데이터 기반 토큰화(Tokenization) 기법 적용
- Ring - Reduce 기반 분산 학습 기술 사용



> 특징

- T-Brain에서 개발
- BERT 의 한국어 성능 한계를 극복하기 위해 개발
- 수백만 개의 한국어 문장으로 이루어진 corpus 를 학습
- PyTorch, TensorFlow, ONNX, MXNet 등 다양한 딥러닝 API를 지원

DEVELOP MODEL

성능 테스트 후 보완

TEST 1

댓글 2000개 [10 epochs] : train acc (0.81) loss (0.73) val acc (0.45) - 과적합

★★★★★ 4

ryun***** · 22.06.22. | 신고
컬러: 연청 / 사이즈: L / 길이: 기본

신축성 좋고 얇아서 시원하네요 대신 컬러가 화면이랑 좀 다른 느낌이예요 연청이 아니라 중청같은

★★★★★ 3

ssin***** · 22.06.25. | 신고
컬러: 연청 / 사이즈: S / 길이: 기본

160인데 기본했어요 이런바진긴게이쁠거같아서요 ㅋ 운동화신어도 다행히 끌리진않네요

>> 리뷰 내용을 입력해주세요 : 신축성 좋고 얇아서 시원하네요 대신 컬러가 화면이랑 좀 다른 느낌이예요 연청이 아니라 중청같은 3점 입니다.

>> 리뷰 내용을 입력해주세요 : 160인데 기본했어요 이런바진긴게이쁠거같아서요 ㅋ 운동화신어도 다행히 끌리진 않네요 3점 입니다.

TEST 2

댓글 50000개 [10 epochs] : train acc (0.95) loss (0.06) val acc (0.72) - 과적합 / 성능 개선 결과

★★★★★ 4

ryun***** · 22.06.22. | 신고
컬러: 연청 / 사이즈: L / 길이: 기본

신축성 좋고 얇아서 시원하네요 대신 컬러가 화면이랑 좀 다른 느낌이예요 연청이 아니라 중청같은

★★★★★ 3

ssin***** · 22.06.25. | 신고
컬러: 연청 / 사이즈: S / 길이: 기본

160인데 기본했어요 이런바진긴게이쁠거같아서요 ㅋ 운동화신어도 다행히 끌리진않네요

>> 리뷰 내용을 입력해주세요 : 신축성 좋고 얇아서 시원하네요 대신 컬러가 화면이랑 좀 다른 느낌이예요 연청이 아니라 중청같은 4점 입니다.

>> 리뷰 내용을 입력해주세요 : 160인데 기본했어요 이런바진긴게이쁠거같아서요 ㅋ 운동화신어도 다행히 끌리진 않네요 4점 입니다.

RESULT

성능 테스트 후 보완

Result 땃글 20만개 [10 epochs] : train acc (0.96) loss (0.1) val acc (0.81)

>> 리뷰 내용을 입력해주세요 : 다음에도 주문할게요
3점 입니다.

>> 리뷰 내용을 입력해주세요 : 이 옷은 사랑입니다
3점 입니다.

>> 리뷰 내용을 입력해주세요 : 너무너무 예뻐서 다음에도 또 주문할 계획이에요
5점 입니다.

>> 리뷰 내용을 입력해주세요 : 옷은 좋은데 배송이 느려요
1점 입니다.

>> 리뷰 내용을 입력해주세요 : 판매자가 너무 불친절합니다
1점 입니다.

>> 리뷰 내용을 입력해주세요 : 배송은 좋은데 옷이 별로예요
1점 입니다.

>> 리뷰 내용을 입력해주세요 : 이상하다
1점 입니다.

>> 리뷰 내용을 입력해주세요 : 뭐가 좋은가요?
3점 입니다.

TEAM MEMBER & ROLE

| 윤소미

자료수집

데이터 핸들링

논문 분석

사전 생성

모델 구현

| 장주찬

자료수집

데이터 크롤링

TEXT Augmentation

토픽 모델링

시각화

| 전규원

자료수집

데이터 크롤링

TEXT Augmentation

사전 생성

SAMPLING

PPT

| 한정현

자료수집

논문 분석

TEXT Augmentation

토픽 모델링

SAMPLING

| 허 권

자료수집

데이터 핸들링

논문 분석

TEXT Augmentation

모델 구현

FIN.

CODEnter

장주찬 | 한정현 | 허권 | 윤소미 | 전규원

APPENDIX

KoBERT

한국어 언어 모델 : Korean Pre-trained Language Models

Model	Accuracy
BERT base multilingual cased	0.875
KoBERT	0.901
KoGPT2	0.899

< Naver Sentiment Analysis >

APPENDIX

KoBERT — 모델 사이즈 대비 성능이 높다.

< Encoder Only Model (BERT LINE) >

Model	Model Size (vocab)	Training Corpus	Detail About Model	
✓ KoBERT [SKT]	8,002	Wikipedia 5M Sents 54M Words	[텍스트 분류] NSMC Acc 89.63	[개체명 인식] Naver-NER F1 86.11
KR-BERT - Character - sub-character	[Character] [sub-character] 16,424 12,367	2.47GB 50M Sents 233M Words	-	
KorBERT - Morphology - WordPiece	[Morphology] [WordPiece] 30,349 30,797	기사, 백과사전 23GB 4.7B Morphs	- ETPI 형태소 분석 API 사용 [Morphology] [WordPiece] [기계독해] - KorQuAD : 1.0 - EM : 86.40% - F1 : 94.18%	 [의미역 결정] Korean Propbank F1 85.77% 85.10% - KorQuAD : 1.0 - EM : 80.70% - F1 : 91.94%