



Ecommerce Clothing Reviews

Jung Kang MIN



Project Overview

언택트 시대가 도래하면서 의류, 식료품, 전자제품 등 다양한 상품을 인터넷을 통해 거래하는 상황

- 수많은 상품이 인터넷상에서 전시되어 있지만, 실제로 사용한 상품이 아니라면 나에게 맞는 상품인지 알기어려움
- 이러한 단점극복을 위해서 이커머스 웹사이트에서 상품 리뷰 글을 남길 수 있도록 하고 최대한 많은 리뷰가 나올 수 있도록 장려
- 소비자는 이러한 리뷰를 통하여 실제 후기와 상품 정보에서 미처 파악하지 못한 정보들을 얻을 수 있으면 긍정적인 리뷰가 많은 상품에는 신뢰를 가진다
- 판매자 입장에서 리뷰가 많은 상품을 마케팅 요소로 사용할 수 있으며 소비자의 의견을 반영할 수 있는 장점이 있다
- 따라서 리뷰는 매우 중요한 특성 데이터!!

Well, the review is
important.

well..



Table of contents

01

Data

EDA

02

BaseModel

DesicionTree

03

Analysis

Timeline

04

Conclusion

Clothing ID	Age	Title		Review Text	Rating	Recommended IND	Positive Feedback Count	Division Name	Department Name	Class Name
767	33	NaN	Absolutely wonderful - silky and sexy and comfy		4	0	0	Initmates	Intimate	Intimates
1080	34	NaN	Love this dress! it's sooo pretty. I happene...		5	0	4	General	Dresses	Dresses
1077	60	Some major design flaws	I had such high hopes for this dress and reall...		3	1	0	General	Dresses	Dresses
1049	50	My favorite buy!	I love, love, love this jumpsuit. it's fun, fl...		5	0	0	General Petite	Bottoms	Pants
847	47	Flattering shirt	This shirt is very flattering to all due to th...		5	0	6	General	Tops	Blouses

Title WordCloud

Do you see?

- Love
- Beautiful
- Great
- Perfect
- Top



Review WordCloud

Do you see?

- Top
- Dress
- Look
- Shirt
- Fabric



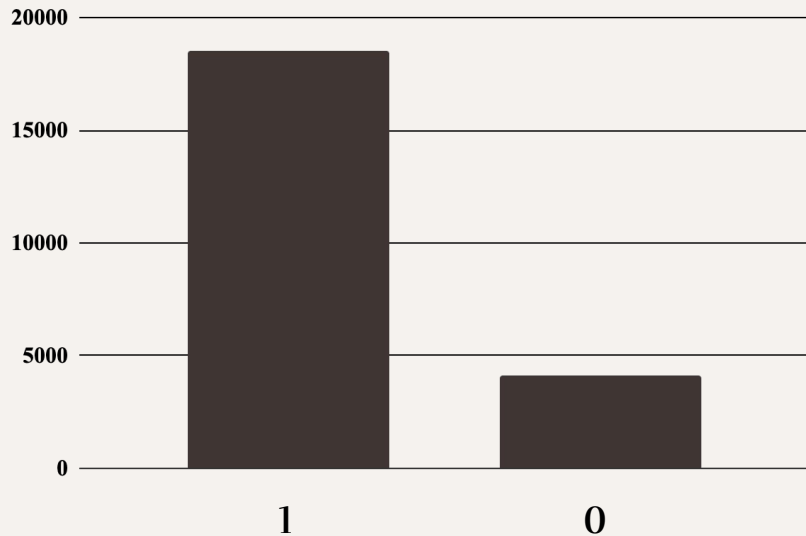
This is Recommended IND

0

Represents
recommendation

1

Represents
Non - recommendation



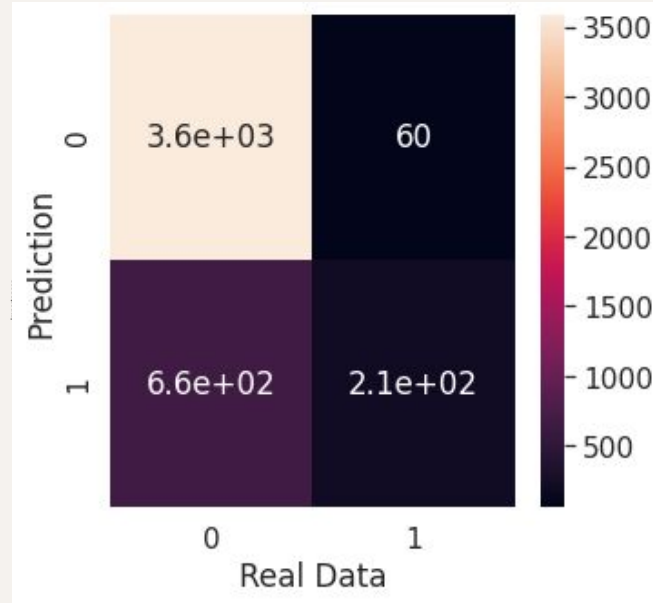
1 represents recommendation,
and 0 represents non-recommendation

DecisionTree (base model)

3592 , 60
665 , 212

Recall score:
0.24173318129988597

Precision score:
0.7794117647058824



0 번째 테스트 데이터 문장: Absolutely wonderful - silky and sexy and comfortable

0 번째 테스트 데이터의 예측 결과: 0, 실제 데이터: 0

1 번째 테스트 데이터 문장: Love this dress! it's sooo pretty. i happened to find it in a store, and i'm glad i did bc i never would have ordered it online bc it's petite. i bought a petite and am 5'8". i love the length on me- hits just a little below the knee. would definitely be a true midi on someone who is truly petite.

1 번째 테스트 데이터의 예측 결과: 0, 실제 데이터: 0

2 번째 테스트 데이터 문장: I had such high hopes for this dress and really wanted it to work for me. i initially ordered the petite small (my usual size) **but i found this to be outrageously small.** so small in fact that i could not zip it up! i reordered it in petite medium, which was just ok. overall, the top half was comfortable and fit nicely, but the bottom half had a very tight under layer and several somewhat cheap (net) over layers. imo, a major design flaw was the net over layer sewn directly into the zipper - it c

2 번째 테스트 데이터의 예측 결과: 0, 실제 데이터: 1

3 번째 테스트 데이터 문장: **I love, love, love this jumpsuit.** it's fun, flirty, and fabulous! every time i wear it, i get nothing but great compliments!

3 번째 테스트 데이터의 예측 결과: 0, 실제 데이터: 0

4 번째 테스트 데이터 문장: This shirt is very flattering to all due to the adjustable front tie. it is the perfect length to wear with leggings and it is sleeveless so it pairs well with any cardigan. **love this shirt!!!**

4 번째 테스트 데이터의 예측 결과: 0, 실제 데이터: 0

Feature importance

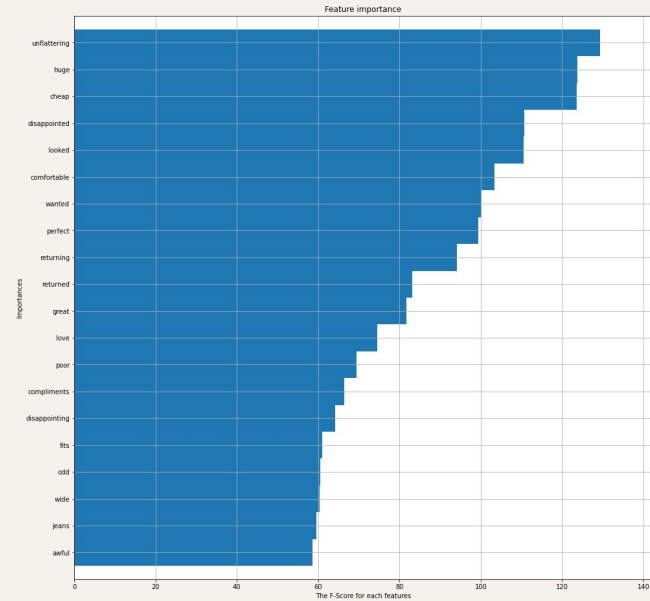
unflattering

huge

cheap

disappoint

look



This is a timeline

DecisionTree

Recall score : 0.24

Precision score: 0.77

GridSearchCV

0.88

01 ——— 02 ——— 03 ——— 04

Data

EDA

KNN	0.88	0.83
NB_M	0.83	0.81
NB_B	0.90	0.86
RF	0.1	0.84
SVM	0.82	0.80
XGB	0.87	0.83

“Artificial intelligence is better than
machine learning.”

conclusion
