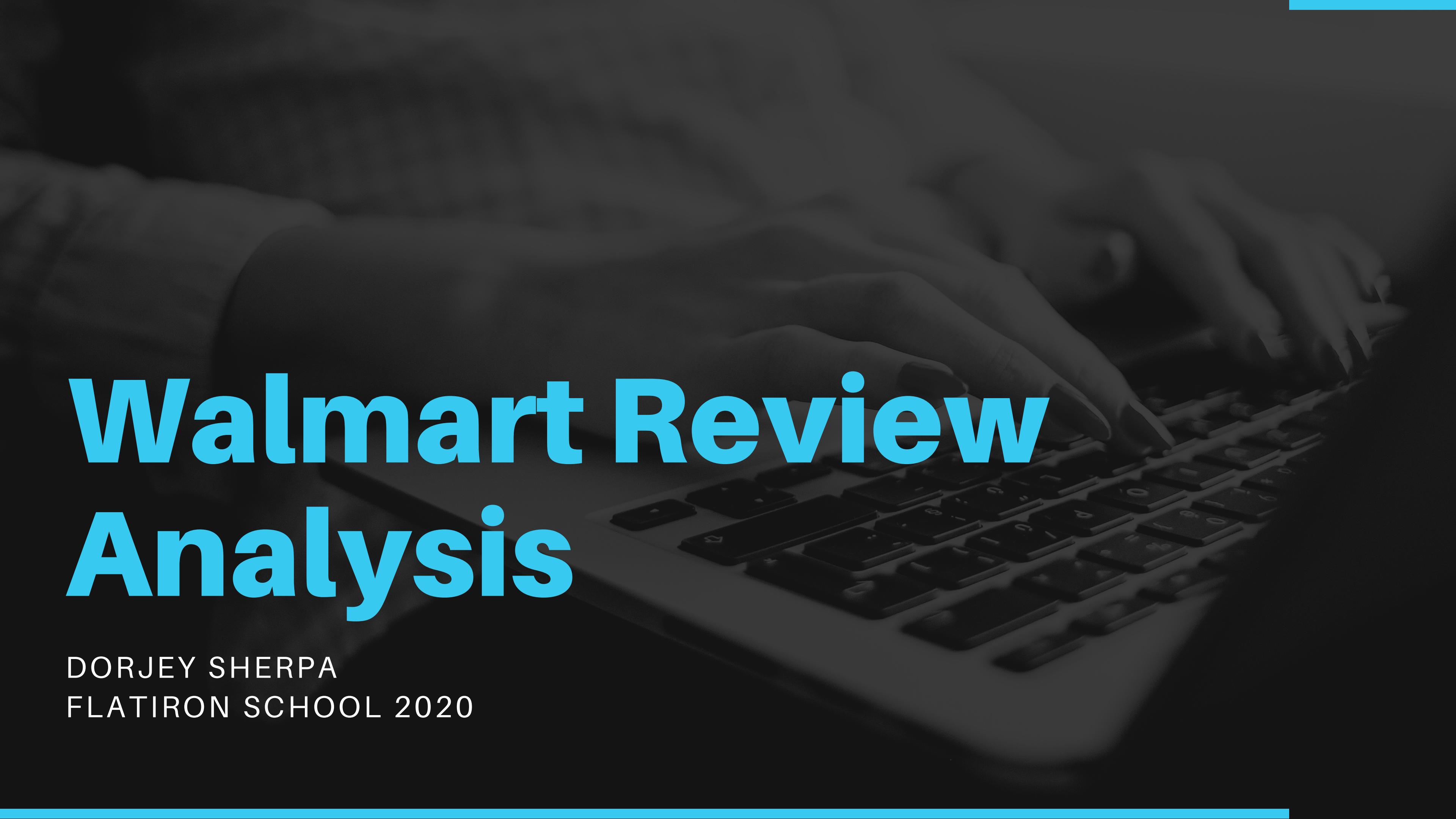
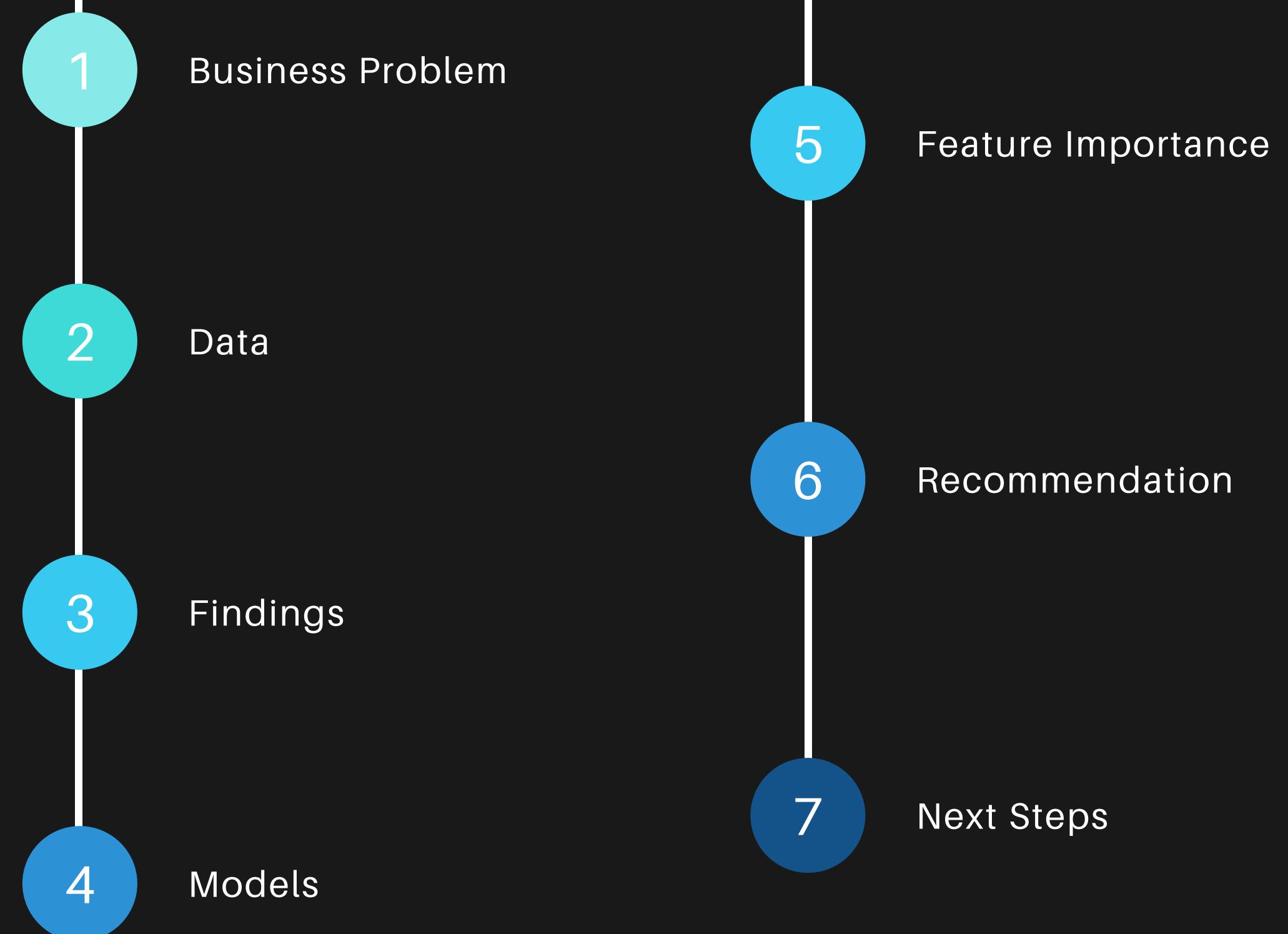


# Walmart Review Analysis



DORJEY SHERPA  
FLATIRON SCHOOL 2020

# Overview

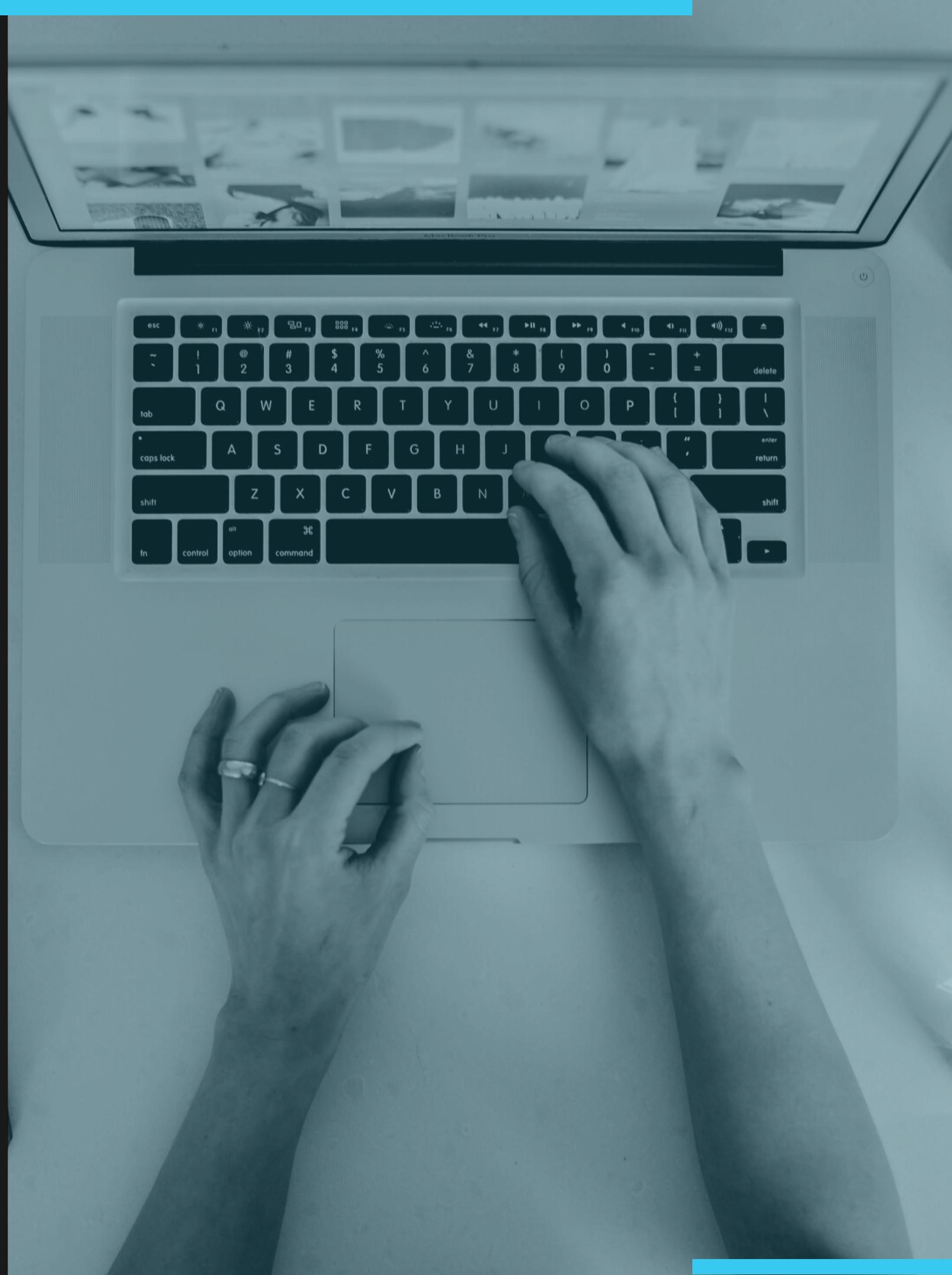


# Business Problem

- Currently: price range and brand.
- Best seller items do not have ratings
- Sense of the online market

## Goal:

- Analyze Walmart Clothing reviews
- Predict whether item is recommend or not by the customers
- Find the features that explain item recommendation



# Data



## DATA GATHERING

scraped directly from Walmart  
webpage



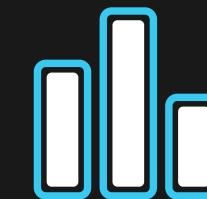
## 4 SECTIONS

Men's clothing  
Women's clothing  
Boys' clothing  
Girls' clothing



## TIMELINE

scraped between November 27 -  
November 30



## OBSERVATIONS

1225 observations with 10  
features

# Data Cleaning, Preprocessing and Feature Engineering

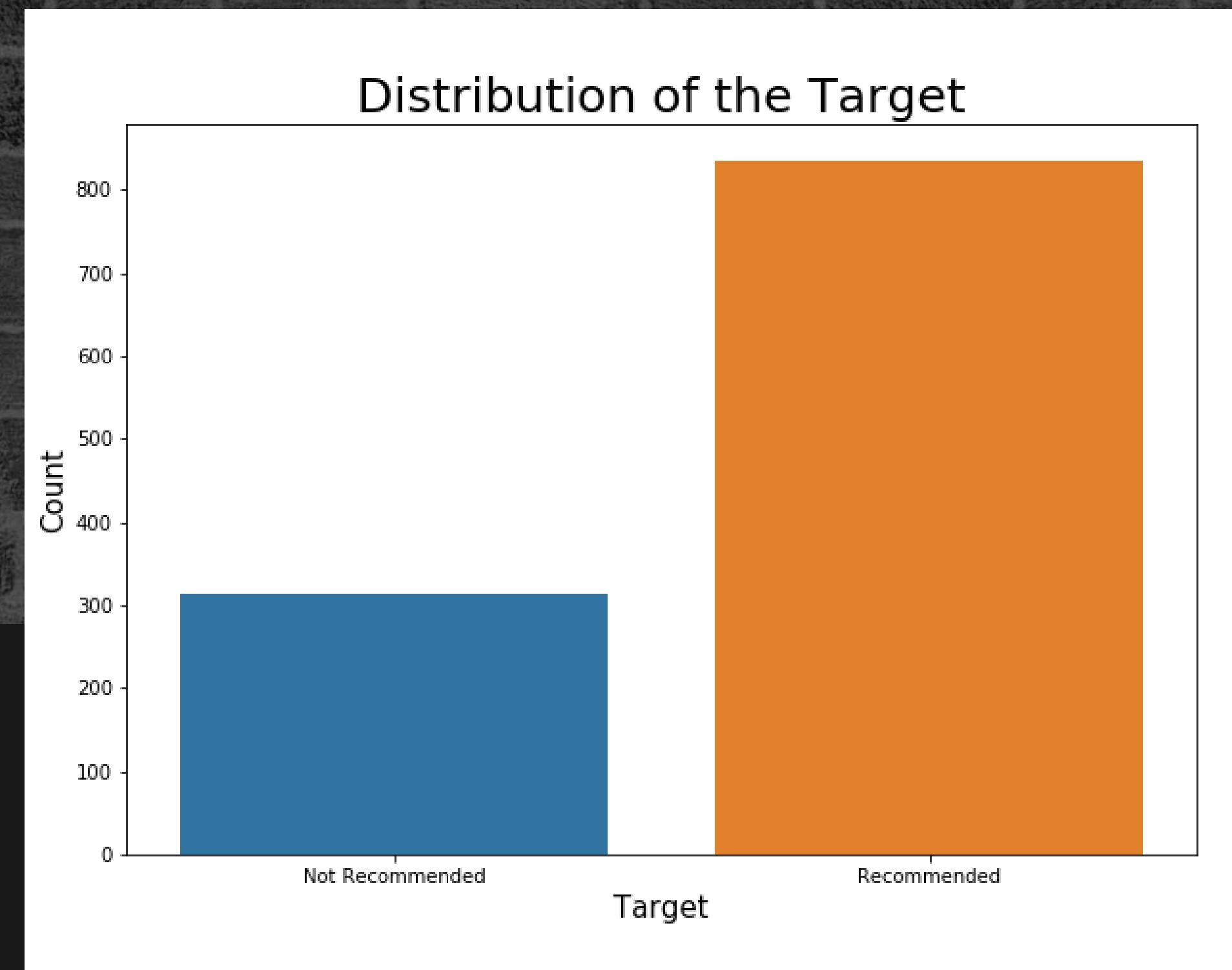
- Clean Continuous Variables
- Regex
- Tokenizer
- Stopwords
- Lemmatizer
- LDA Topic Modeling

## End Result:

- 1149 observations with 29 features

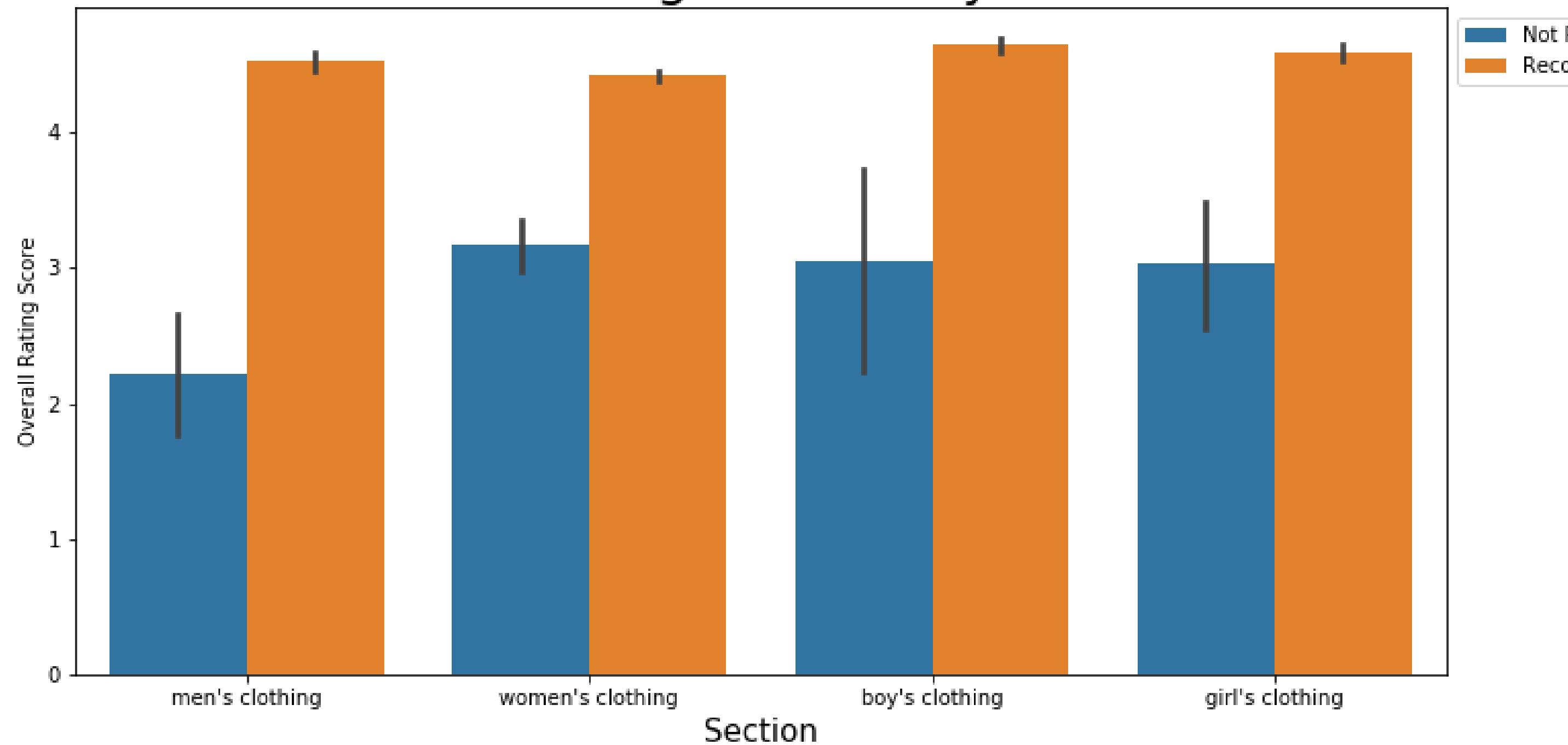
# Findings

Categorize the Recommended Score:  
Above 83% (inclusive) and  
below 83%



# Findings

## Overall Rating of items by Section



Higher the rating, more likely the item is to be recommended

# Modeling

- Logistic Regression
- Bernoulli Naive Bayes
- Support Vector Classifier - Radial Basis Function
- Random Forest Classifier
- GridSearchCV with Random Forest Classifier
- GridSearchCV with Logistic Regression

# Metric

F1 Score

# Findings

Models

F1 Score

GridSearch  
Log Reg

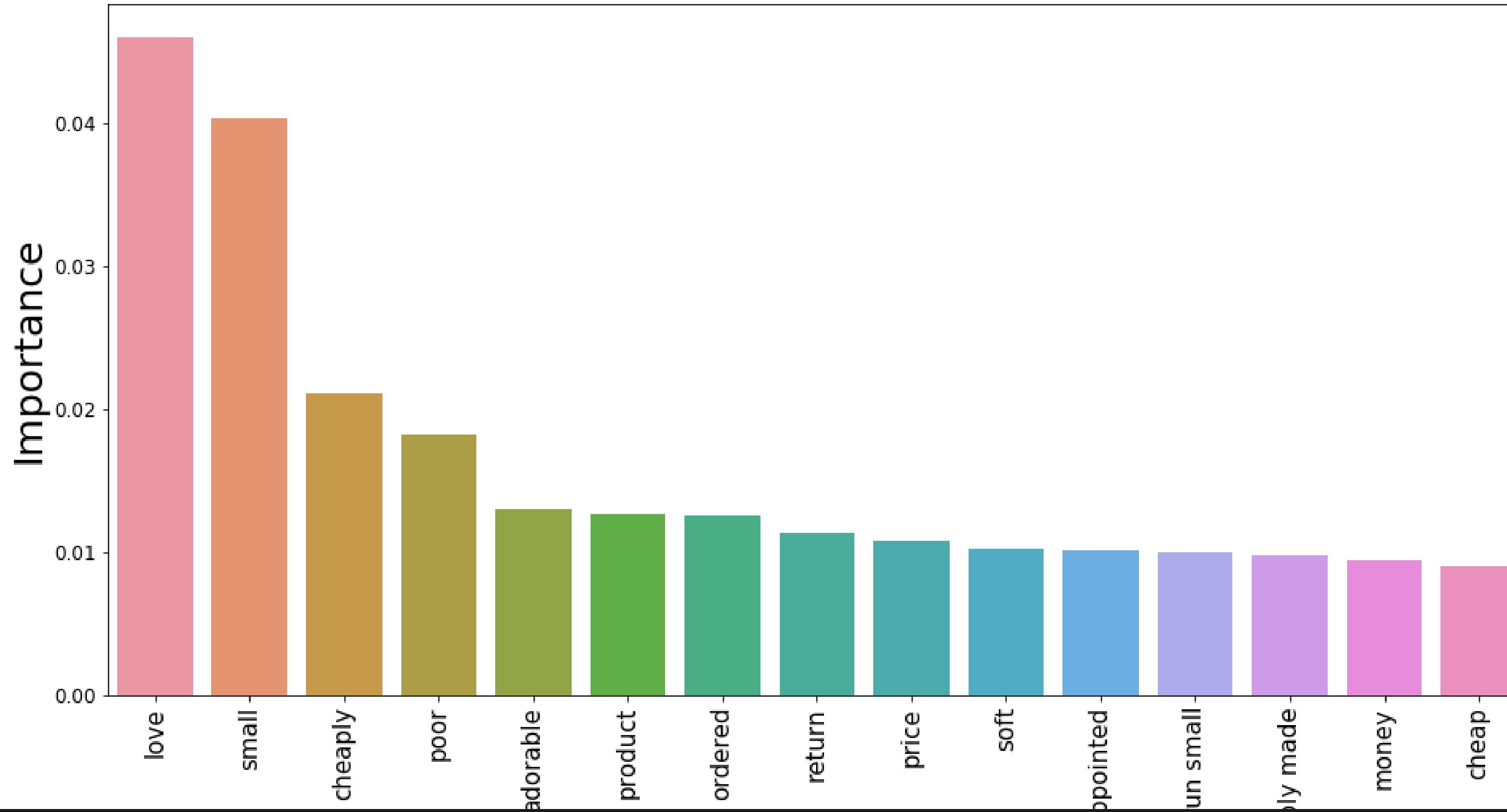
0.839

GridSearch  
Random Forest

0.88

# Feature Importance

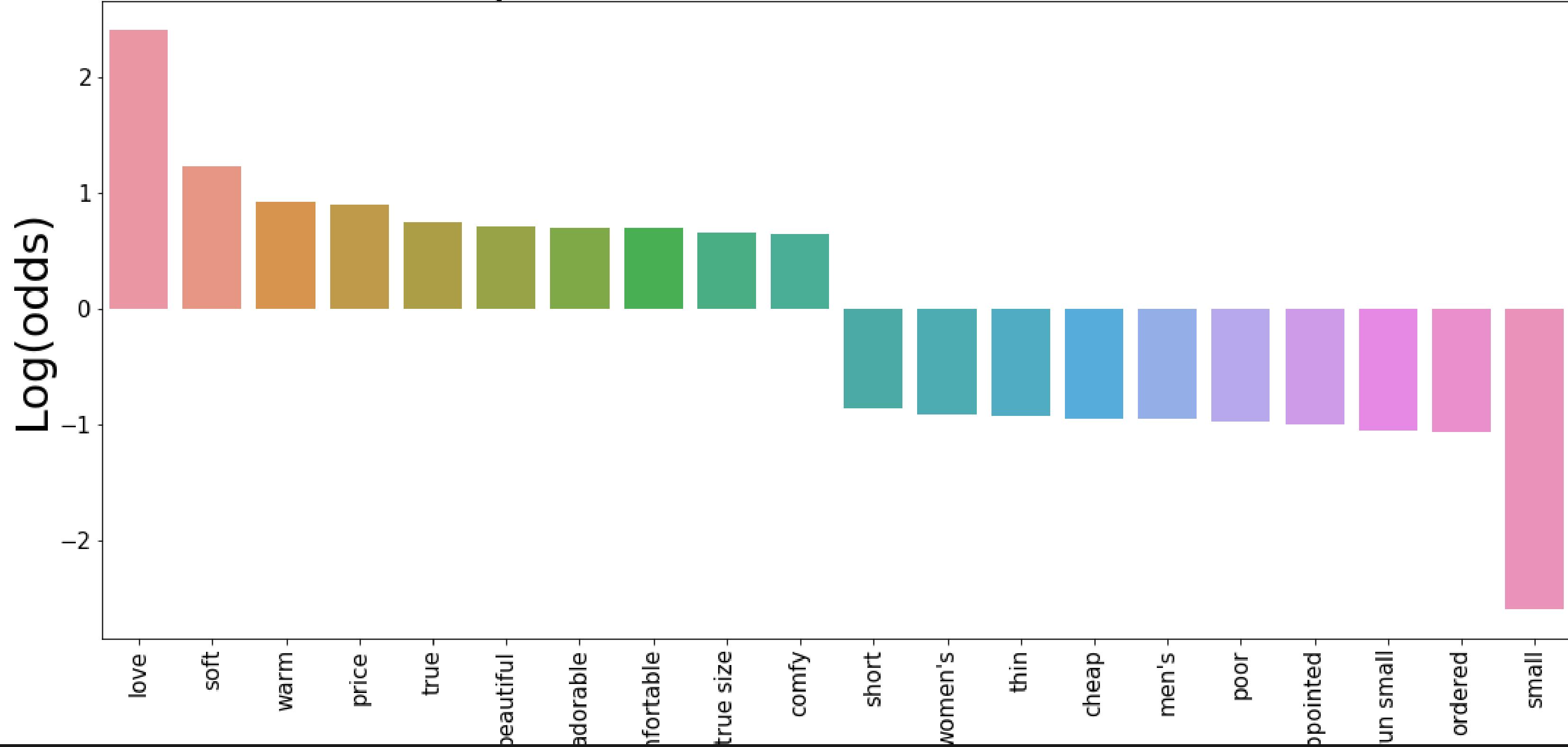
GridSearchCV with RandomForestClassifier



# Feature Importance

## Logistic Regression

### Top 10 and Bottom 10 Features



# Recommendation

## Fit importance

*"My niece loved the bright pink color  
and it fit her perfectly!"*

*"was disappointed that the sizing did  
not compare to the previous ones I  
purchased from you."*

*"Runs really small. I will be returning  
it..but cute"*

# Recommendation

## Quality Control

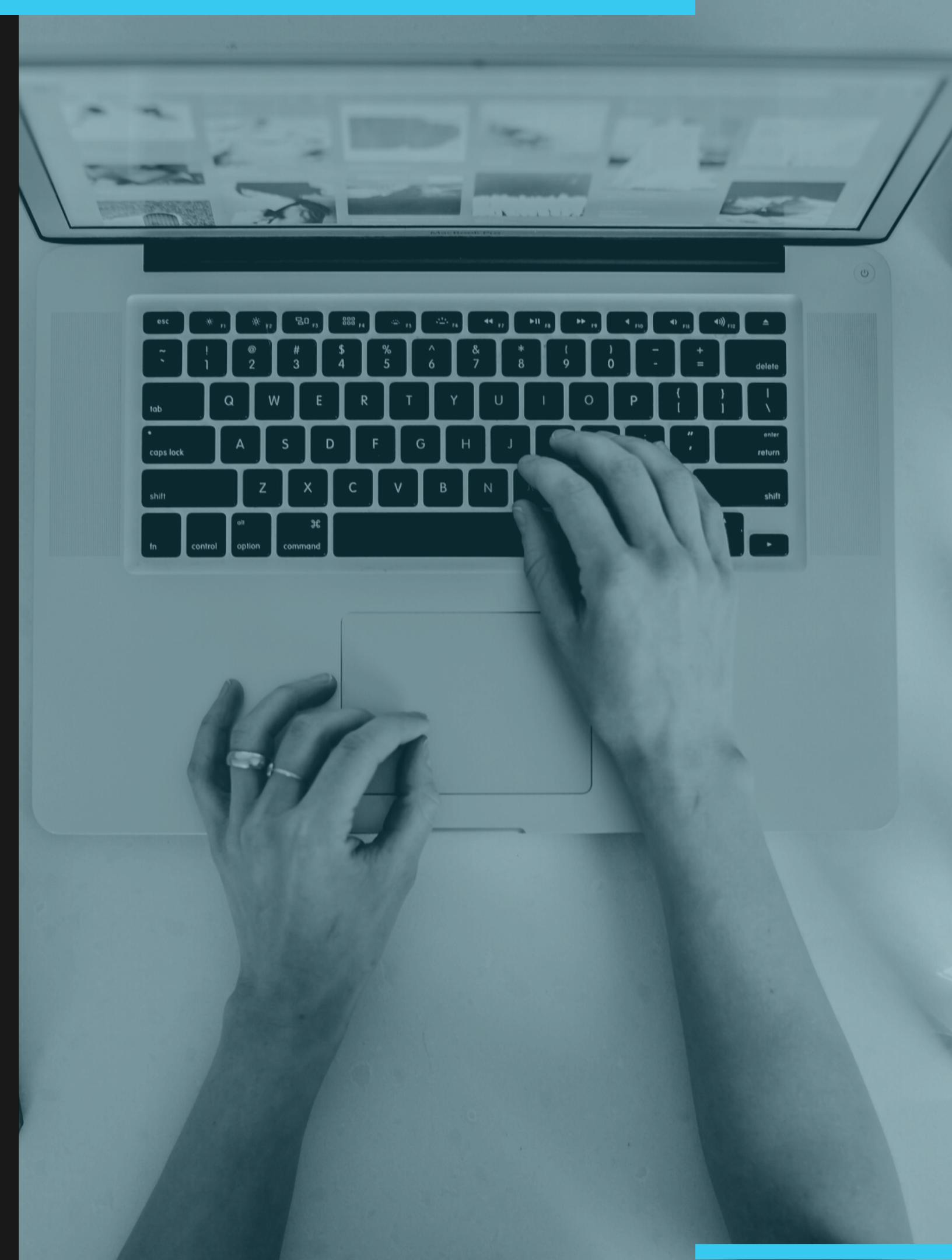
*"When we opened the vacuum sealed packaging that this Mario suit came in, we were completely overwhelmed by a smell of mold and mildew."*

*"These come from overseas so the sizing is way small. I ordered a womens medium anf it was like an extra small. They are super thin material and the imprint is cheaply made."*

*"First, the sweater came smelling really bad like vinegar. Worst of all, the design came off after just one wash!"*

# Next Steps

- Lemmatization vs Stemming vs Nothing
- Compare brands
- Scraping reviews across seasons



## Future Steps

### Slip

*"Did not fit my feet well. **Slipped** off easily"*

*"These **slippers** run SO Small!"*

*"These are the best **slipper** socks I have ever owned!"*

*"Hubby also claims they **slip** down easily if not tied tightly."*

*"I sadly had to return my ballet **slipper** pink because I ordered the wrong size."*

# Thank you

For further clarification or inquiries please  
contact below:



dorjeys3@gmail.com



github.com/dorjeys3



twitter.com/dorjeys3

