FIT PREDICTIONS

minimize returns | maximize revenue
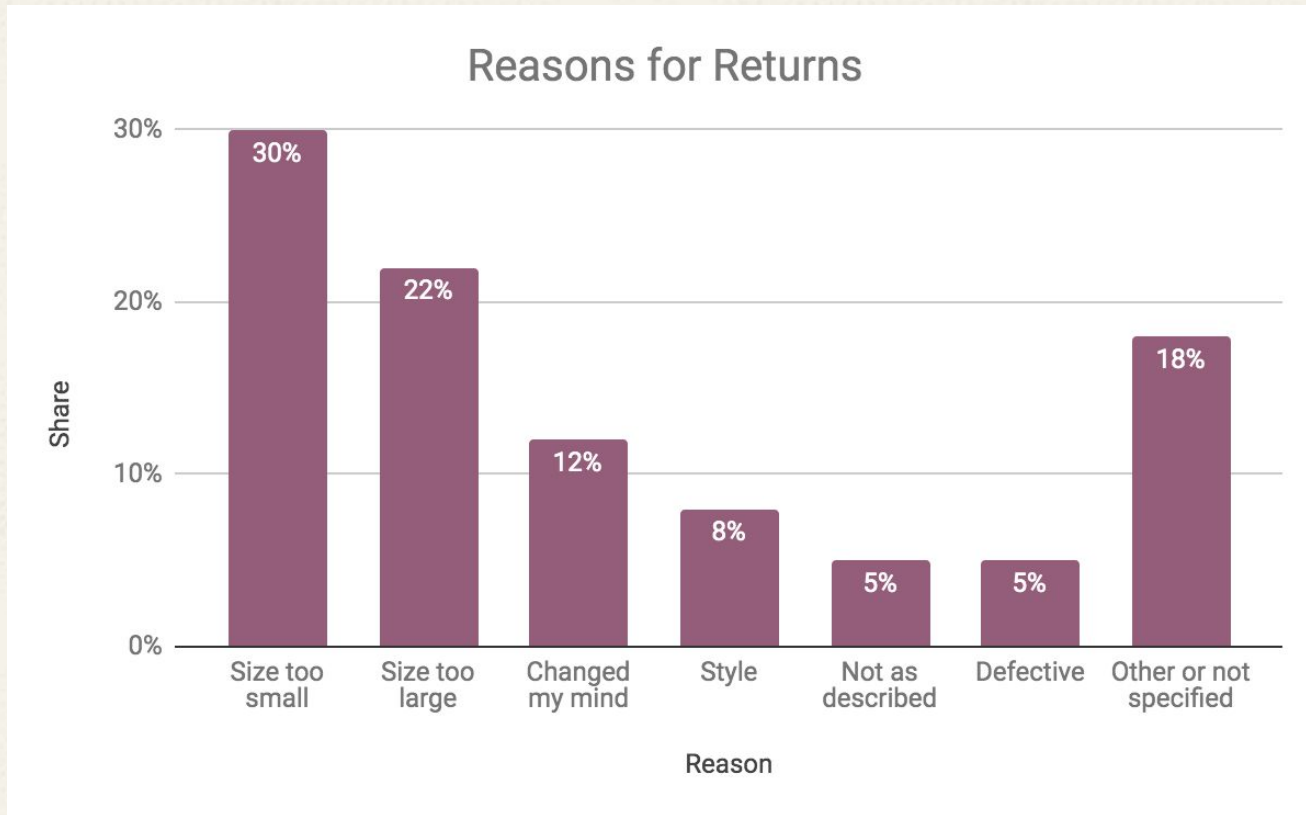
DSI Capstone Project

Arielle Miro

"

Consumer preference-based return reasons (e.g., size, fit, style, etc.) tend to drive around 72% of all returns in fashion product categories..

2

CURRENT RETURNS STATS

Reasons for Returns

Incorrect sizing accounts for over 50% of returns in the e-commerce retail space.

3

1. PREDICT "FIT"

Battle the plague of return rates through smart sizing predictions

DATA FRAMEWORK

Rent the Runway measurements, user info, and review information.

SIZE & CATEGORY

- User
- Item
- Category
- Occasion
- Size
- Body Type
- Weight
- Height
- Age

REVIEWS

- Review Summary
- Full Review

PREDICTIVE MODELS

Multiclass Classification

SIZE & CATEGORY

- All Numeric Columns
- Random Forests
- Best Score: 73.8%

REVIEWS

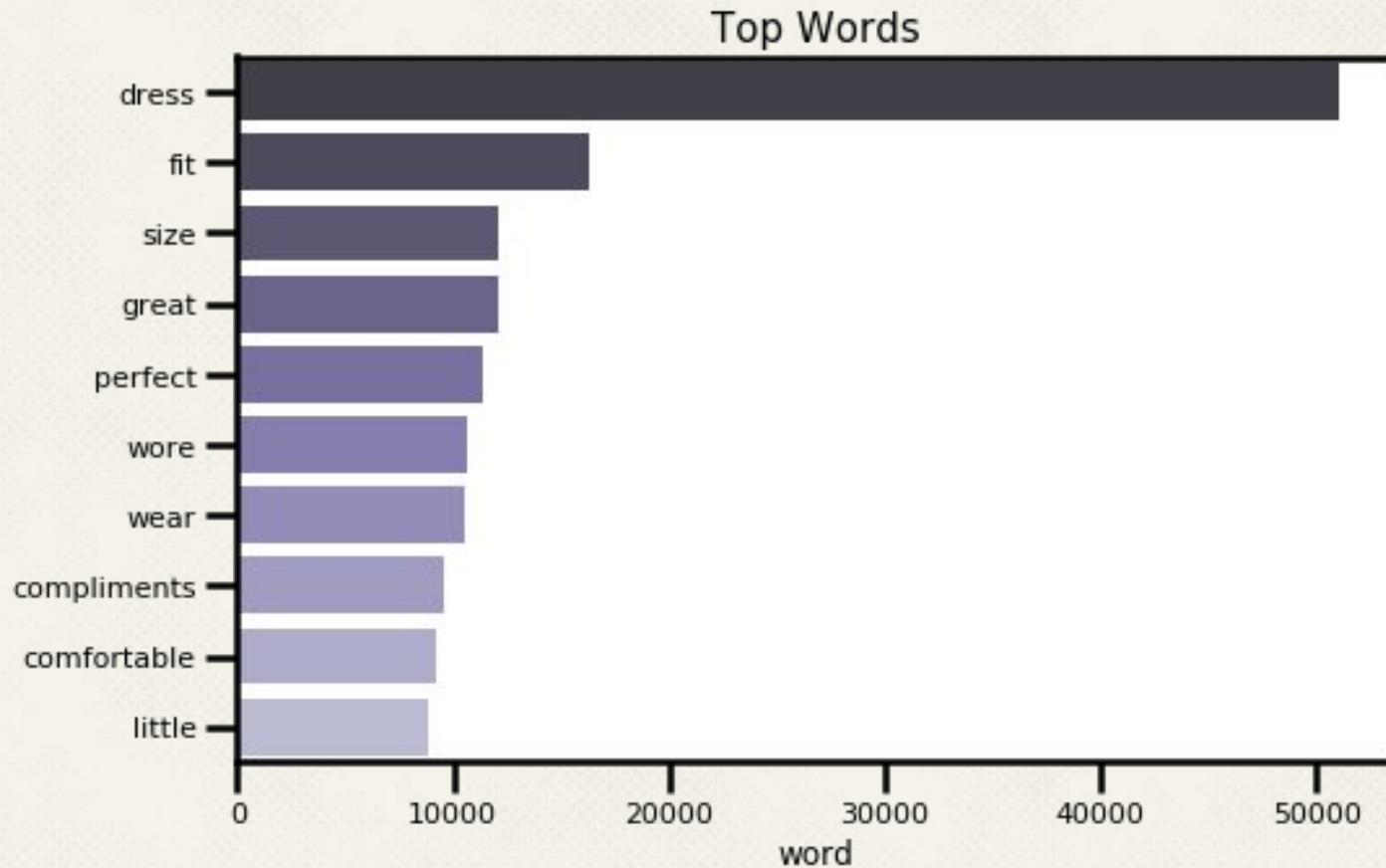- NLP
- CountVectorizer
- Random Forests
- Best Score: 82.7%

CHALLENGES:

Unbalanced Classes , Overfit Models

NLP Classification
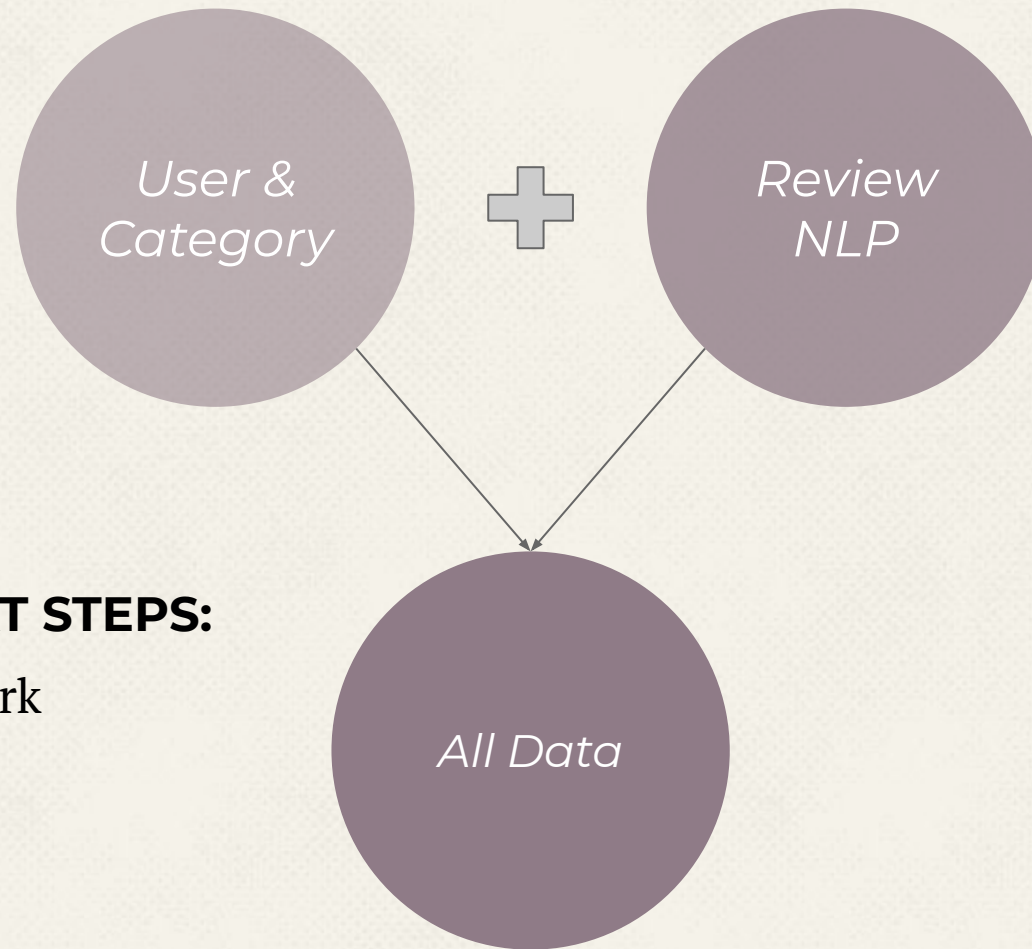

Top Words

MODEL IMPROVEMENTS

User & Category

➕

Review NLP

All Data

POSSIBLE NEXT STEPS:

- Neural Network
- Spacy
- Word2Vec

2. RECOMMEND

Increase revenue & enhance user experience with smart product recommendations

COLLABORATIVE RECOMMENDERS

USER BASED

Recommends based on similar users. Assumes that similar users have similar tastes.

ITEM BASED

Recommends based on similarities between items.

CHALLENGES:

User profiles are based on sizing rather than preference.

Data is anonymized (users & items).

CONTENT RECOMMENDER

- Based on the item reviews
- Using NLP
- Matrix of Vectorized Words

"Beautiful dress and even better in person! "
Item #141688

Top 5 Recommendations for item #**29207**

user_id	age	body_type	category	fit	height	item_id	rating	rented_for	size	weight
46348	44	hourglass	shirtdress	1	5. 4	1840637	6	party	4	108
401375	39	petite	gown	1	5. 2	141688	8	formal affair	12	123
338462	34	petite	gown	1	5. 1	832622	10	formal affair	1	103
831814	32	hourglass	dress	1	5. 10	249458	10	party	24	190
105378	42	straight & narrow	dress	1	5. 3	172027	8	party	1	104
297705	31	hourglass	gown	1	5. 6	1714731	10	wedding	24	190
980350	28	full bust	gown	0	5. 8	149655	8	wedding	17	150
256675	35	athletic	dress	0	5. 6	345146	10	wedding	16	130
763383	32	full bust	dress	1	5. 4	125424	10	other	20	160
644888	30	athletic	sheath	1	5. 9	987743	10	wedding	16	178

"Go up one or two sizes - you'll be happy you did."
Item# 345146

CONCLUSION

💬 NLP PERFORMED "BETTER"

Fit Predictions :

- ○ NLP model reached a higher accuracy score

Product Recommender :

- ○ Using NLP captured user preferences the general size and fit data could not.
- ○ Underlying item similarities were rooted in the text reviews

NEXT STEPS

Build "Master" Model

Leverage AWS to model the combined data for sizing and reviews. Compare this score to the separate models.

Expand on Content Based Recommender

Incorporate sentiment analysis and categorical data.

ADDITIONAL IMPROVEMENTS

Purchase Amount

To perform predictive analytics on revenue projections. Also, calculating revenue loss due to returns.

User Location

Additional profile detail to analyze the consumer demographic beyond age and size.

Item Description

Additional information about the item will allow for more granular predictions. An image of the item may allow for further image recognition comparisons.