Returns Propensity Prediction for Online Orders

RALPH LAUREN

Mrugank Akarte
Kanika Aggarwal
Priyanka Lahoti
Sachit Kumar
Hima Bindu Bhardwaj

Mentor and Industry Affiliates:

Mentor:

Sining Chen, Adjunct Professor at Data Science Institute, Co-Director of Capstones

Ralph Lauren Team:

Rohit Cherian
Nandakumar Sudha
Benjamin Raibaud
Landon He
Danielle Zuckerman
Michael O'Reilly

Business problem:

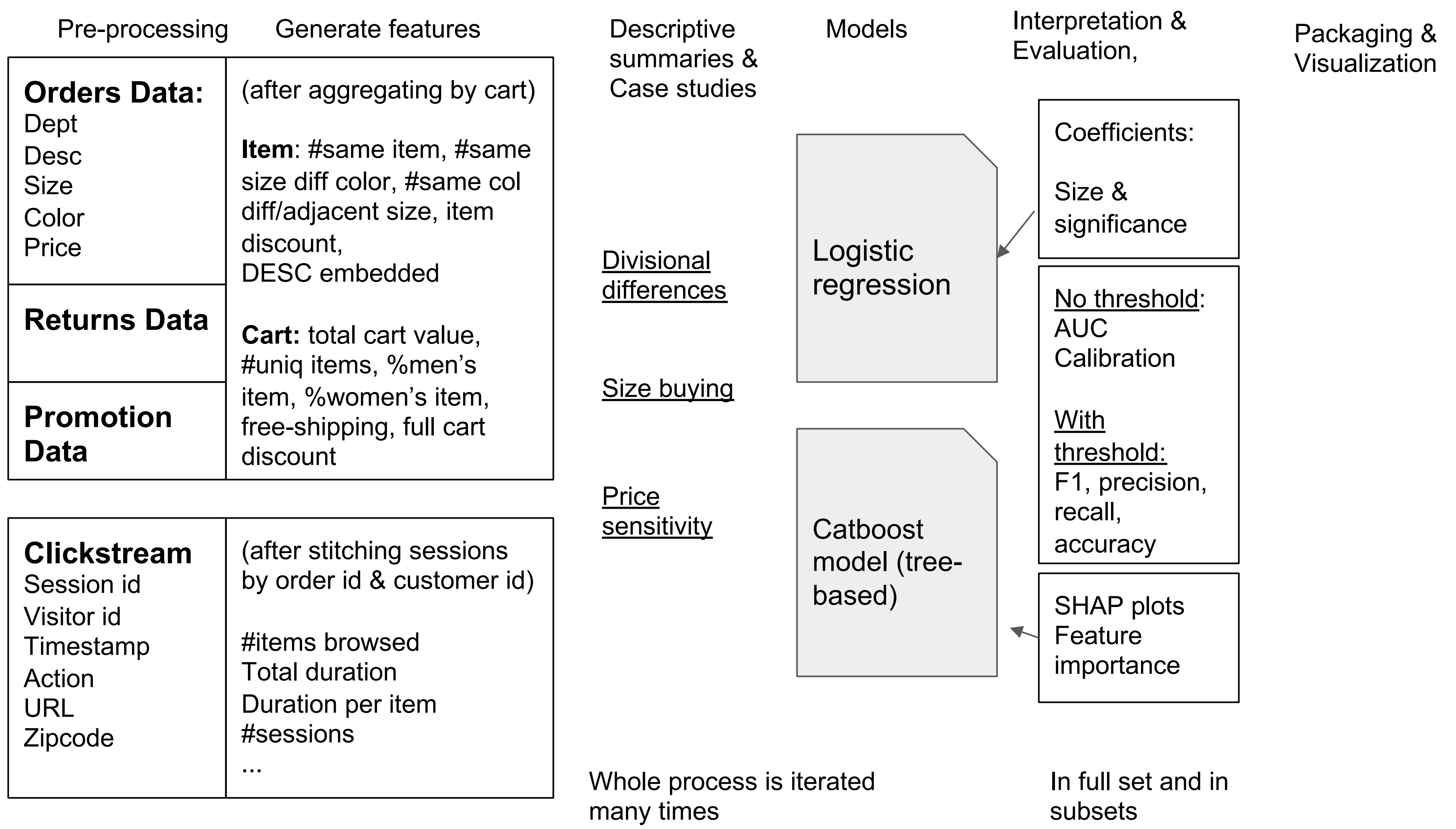
- Since product returns can add a considerable financial, environmental and operational cost factor to a fashion e-commerce business like Ralph Lauren, a returns propensity forecasting model can help allocate personnel for this effort.
- Such study/work can also help Ralph Lauren understand customer purchase behavior which can lead to interventions that can be implemented to reduce returns

In our work, we attempt to understand what factors contribute to the returns of an item and we build a model to predict the return probability of products purchased.

Project Objectives:

To understand factors that lead to a product having high propensity of return.

To develop a model for return predictions. These can be used to provide customized offers for a product with higher probability of return in cart.


Pre-processing
Generate features
Descriptive summaries & Case studies
Models
Interpretation & Evaluation,
Packaging & Visualization
Orders Data:
Dept
Desc
Size
Color
Price
Returns Data
Promotion Data
(after aggregating by cart)
Item: #same item, #same size diff color, #same col diff/adjacent size, item discount,
DESC embedded
Cart: total cart value, #uniq items, %men's item, %women's item, free-shipping, full cart discount
Clickstream
Session id
Visitor id
Timestamp
Action
URL
Zipcode
(after stitching sessions by order id & customer id)
#items browsed
Total duration
Duration per item
#sessions
...
Divisional differences
Size buying
Price sensitivity
Logistic regression
Catboost model (tree-based)
Coefficients:
Size & significance
No threshold:
AUC
Calibration
With threshold:
F1, precision, recall, accuracy
SHAP plots
Feature importance
Whole process is iterated many times
In full set and in subsets

Preprocessing:

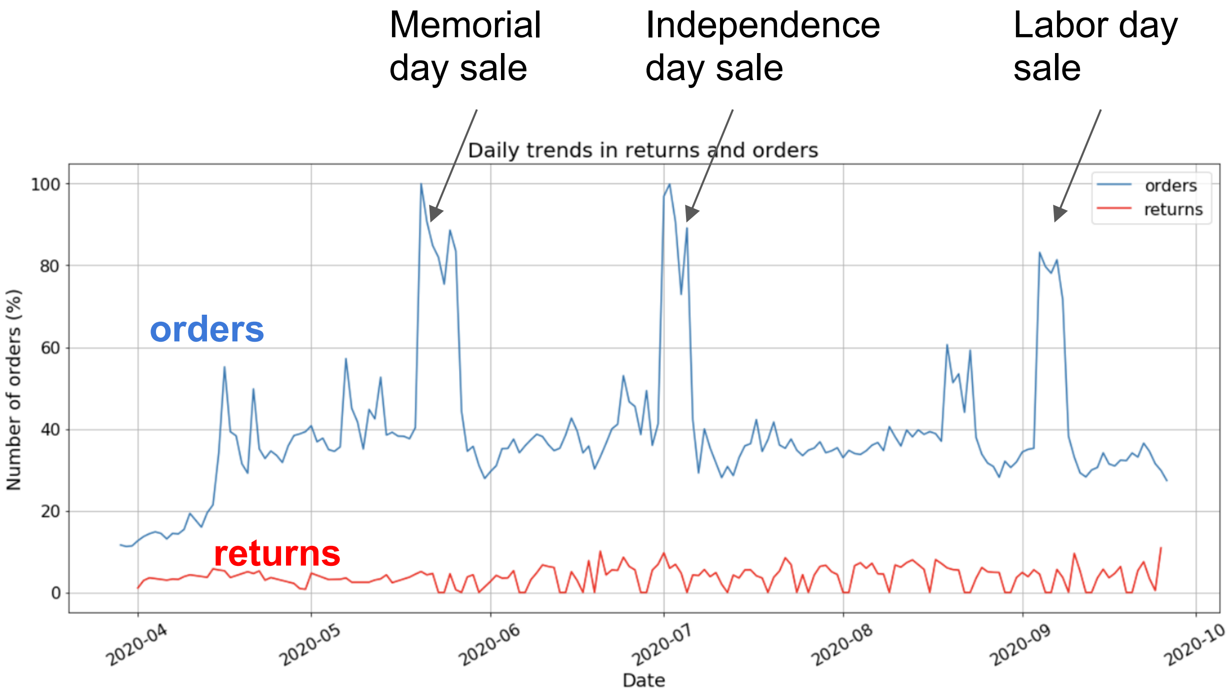
We merged orders, returns and promotion dataset to create a new dataset:

- Initial number of records: 3.84M
- After cleaning: 3.81M

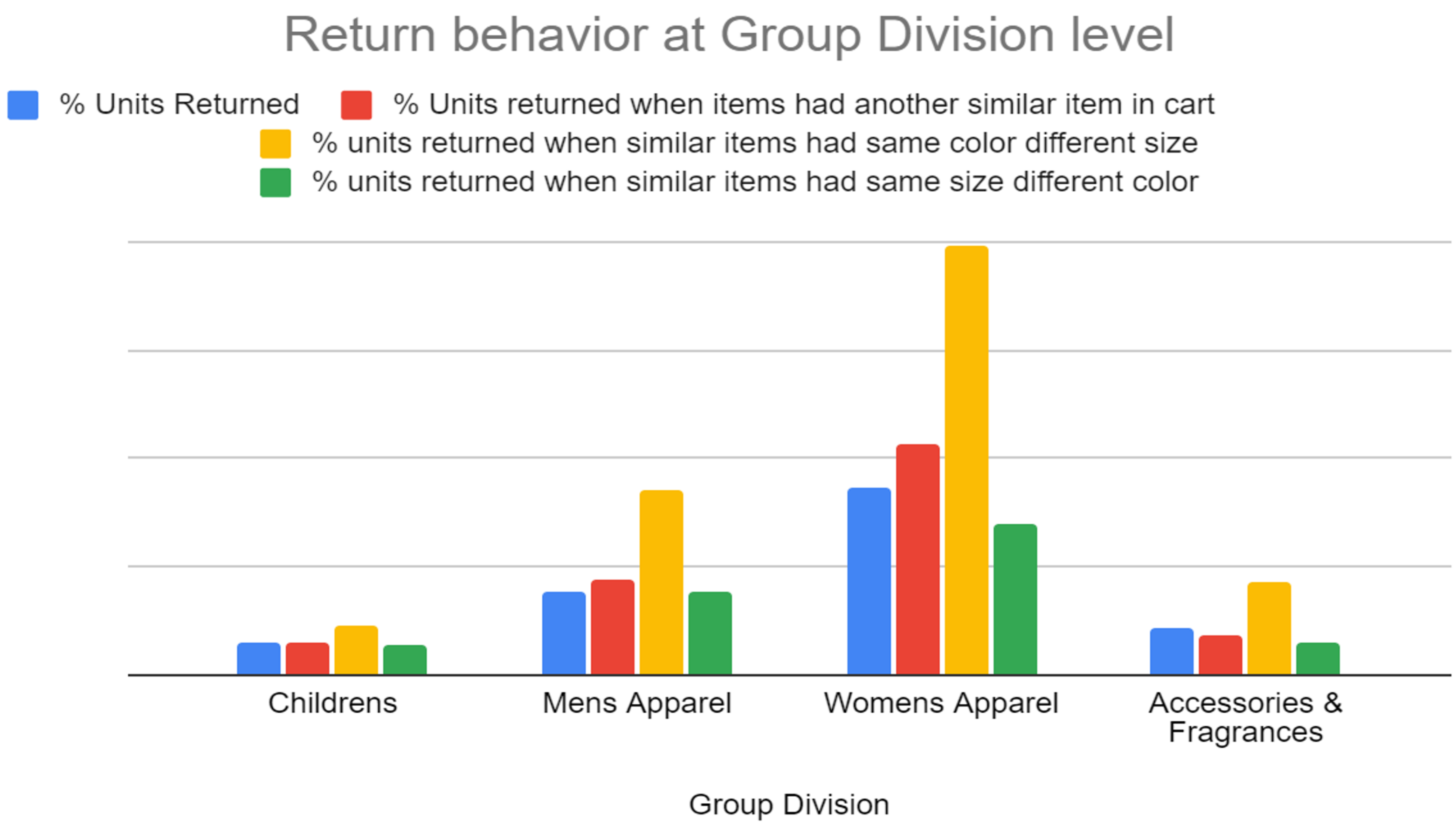
Preprocessing steps:

- Changed date fields to date time format
- Removed records with unknown product details
- Removed orders that had returns but no sales associated with them

Daily trends of orders vs returns:



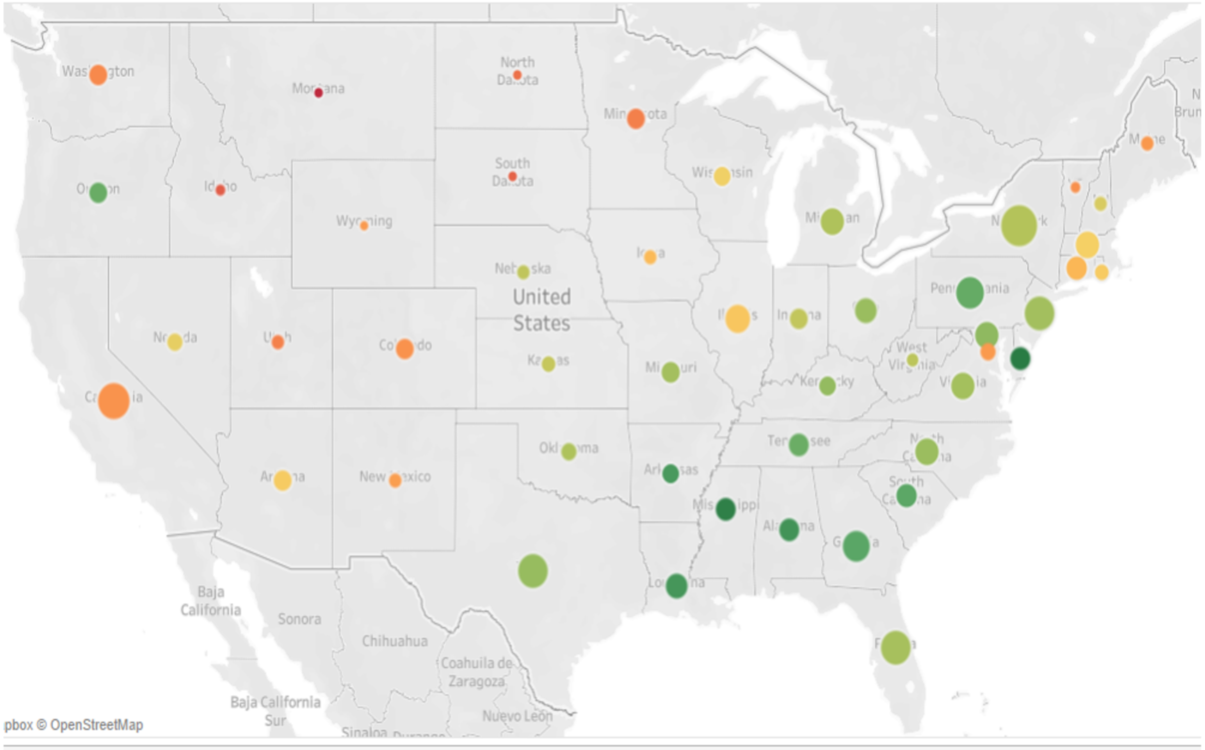
We see a big spike for orders in mid May, early July and in early September. We believe that mid-May spike corresponds to Memorial Day sales, early July spike corresponds to July 4 sales and early September spike corresponds to Labor Day sales.

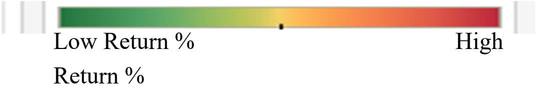


* similar items are those with same style but differ in size or color

State level return % - All departments

State mapped using zip codes from clickstream data





States with higher return% usually has low orders

States like NY, California, Florida show high orders and above average return %

States are represented by dots, the size of the dots are proportional to the sqrt of number of items ordered from that area

color corresponds to the return% - green is low while red is high return %

Models

Logistic Regression

- Baseline model
- Highly interpretable but suffers in performance
- Limited to few categories for categorical columns

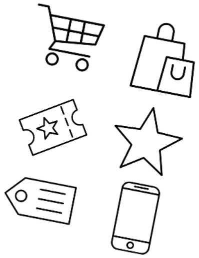
Catboost Classifier

- Gradient boosting library for decision trees
- Can use non-numeric factors for training
- Less interpretable but give better performance
- Shap plots used for model interpretation

Decision Tree
Random Forest
Neural Networks

Experiments with Features

Product Level Features
- Product attributes like price, color, size, etc
- Discounts, coupons
- Log transformations of variables



Order Level Features
- Time related features
- Total cart value
- Unique products in cart
- Similar items in cart
- zipcode

Clickstream features
- Page views, items added to cart, session time
- Session data was not representative of real browsing session, hence these features were discarded

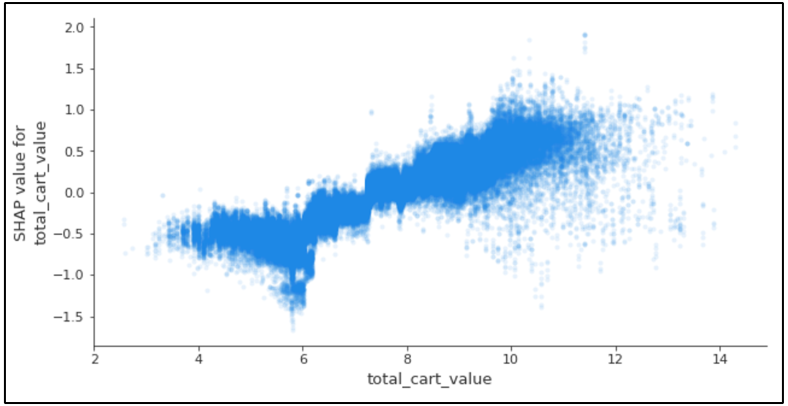
Word Embeddings for product description

- **Model Metrics: AUC, precision and recall**

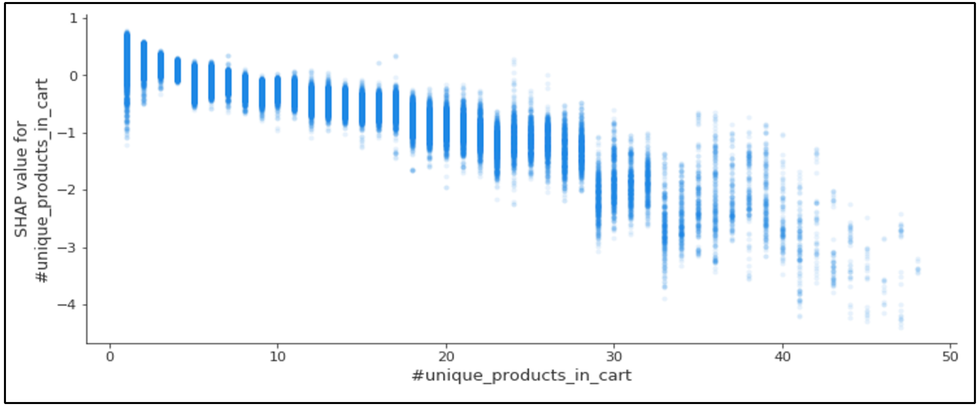
AUC	All department data	Mens Data	Womens Department	Children Department
Logistic Regression	0.704	0.593	0.675	0.642
Catboost Classifier	0.815	0.754	0.721	0.824

- Cart level features like total cart value, number of unique products in cart had higher importance.
- Product level features like product price, item description and department were major driving factors.

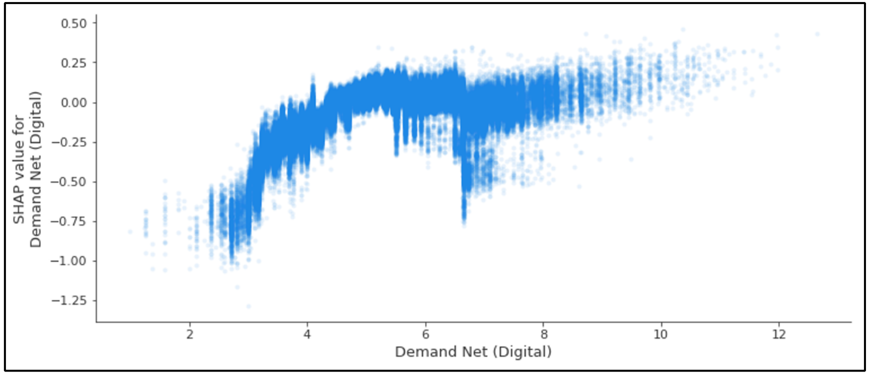
Model Interpretation using SHAP values (log-odds)



Total cart value: In general, return probability increases as total cart price increases.

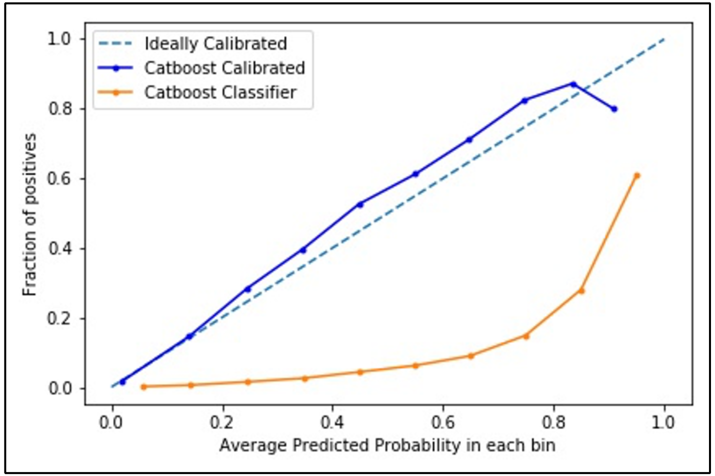


Unique products in cart: The log-odds generally decrease as the number of unique products in cart increases.



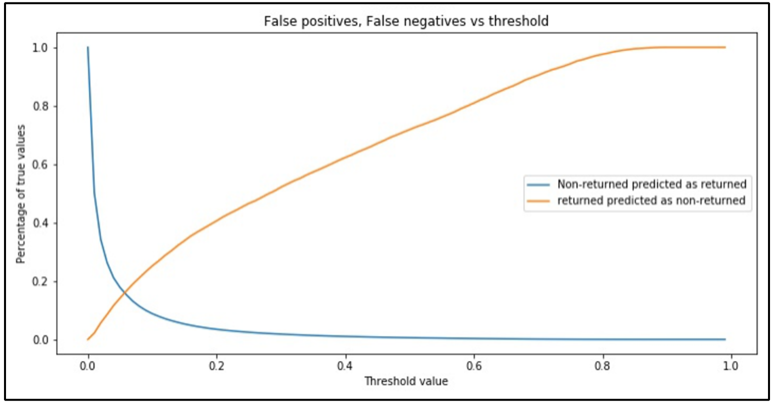
Demand net: Cheaper items are less likely to be returned. Most of products from Children department are cheap as compared to Mens and Women department

Catboost Classifier - Calibration



Outputs of Logistic Regression are already calibrated. The original Catboost Classifier output overpredicts probabilities. Calibrated probabilities can be directly interpreted as confidence level.

This plot on right shows changes in false positives and false negatives with respect to different threshold values.



Learnings

- We observed different return behavior in different departments (Mens, Womens, Childrens)
- Since the model is built only product attributes, it is difficult to capture customer behaviour from a cart.
- Error analysis helps to understand the shortcomings of the model and can help in defining new features for the model.
- Customer attributes such as clickstream data (that represents the whole shopping experience) can be very helpful for a more precise model.

Future Work

- Expand model to include customer specific information
- Explore different word embeddings for other category variables
- Extensive error analysis to understand the unusual patterns in data