

Building a model to predict e-commerce order cancellations

Natalia Shcheglova, Aug 27, 2024

The Problem

- Loss of revenue
- Disruption of fulfillment process
- Costs associated with processing cancelled orders



Why Would We Want To Predict Cancellations?

Reduce cost



Understand customers



Improve operational efficiency



Win back revenue



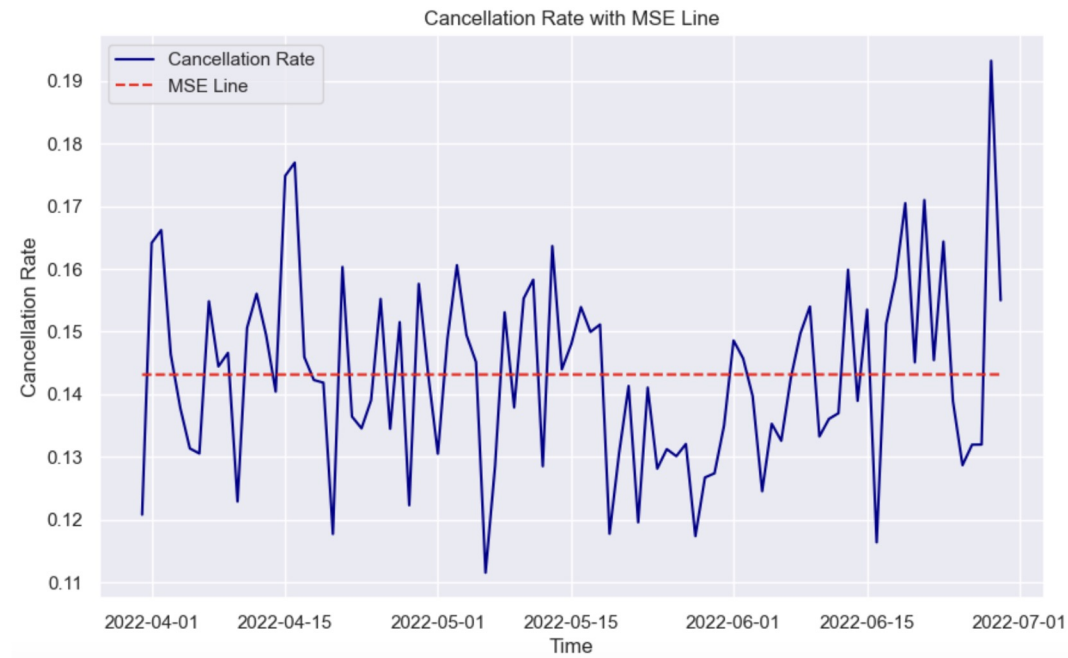
The Data

Amazon Sale Report for India market for Q2, 2022

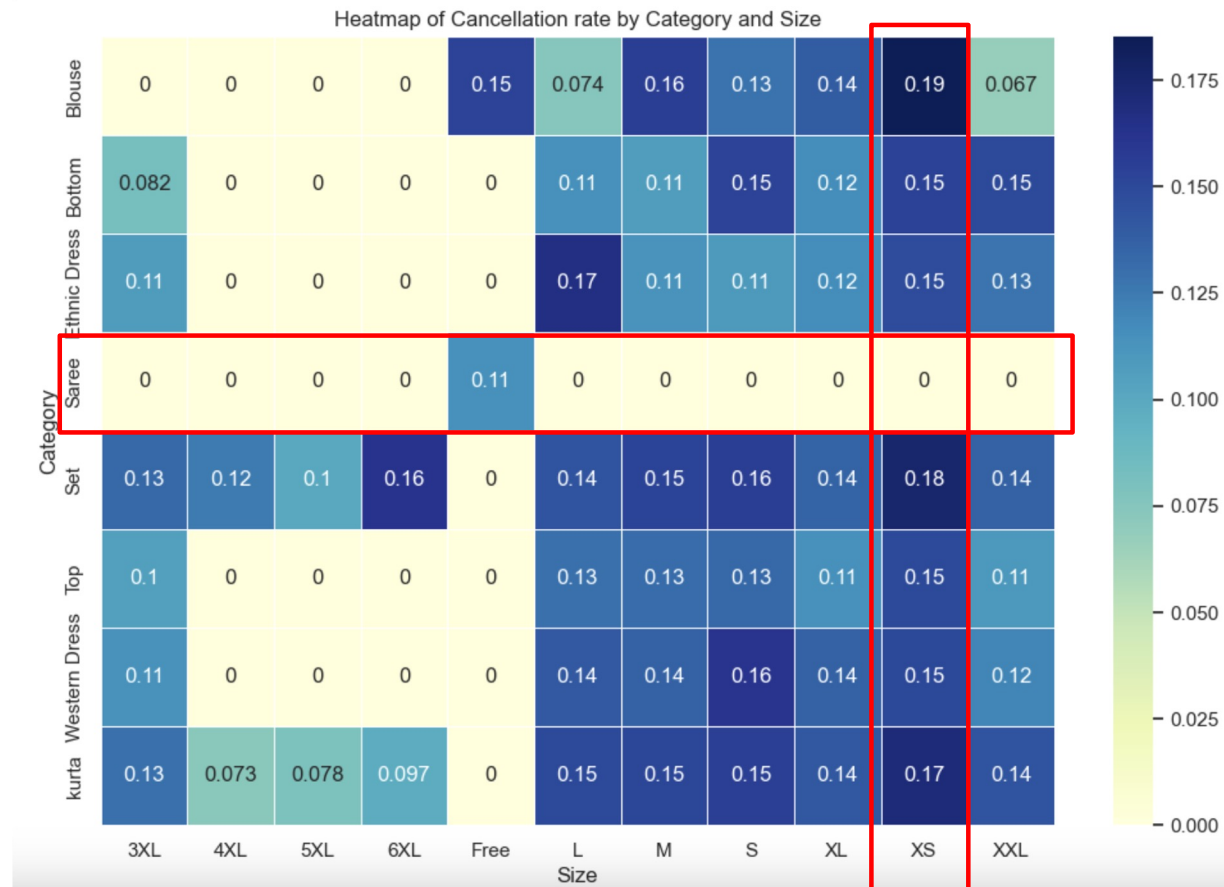
Date	Product details	Price	Shipment address	Shipment details	Promotion
Month	Category		State	Shipment type	Free shipping
Day	Style		City		Free financing
Time of the month	Size		Zip	Fulfilment merchant	Coupon
Weekend	SKU				No promotion

Data Exploration

$$\text{Cancellation rate} = \frac{\text{Number of cancellations}}{\text{Total orders}}$$

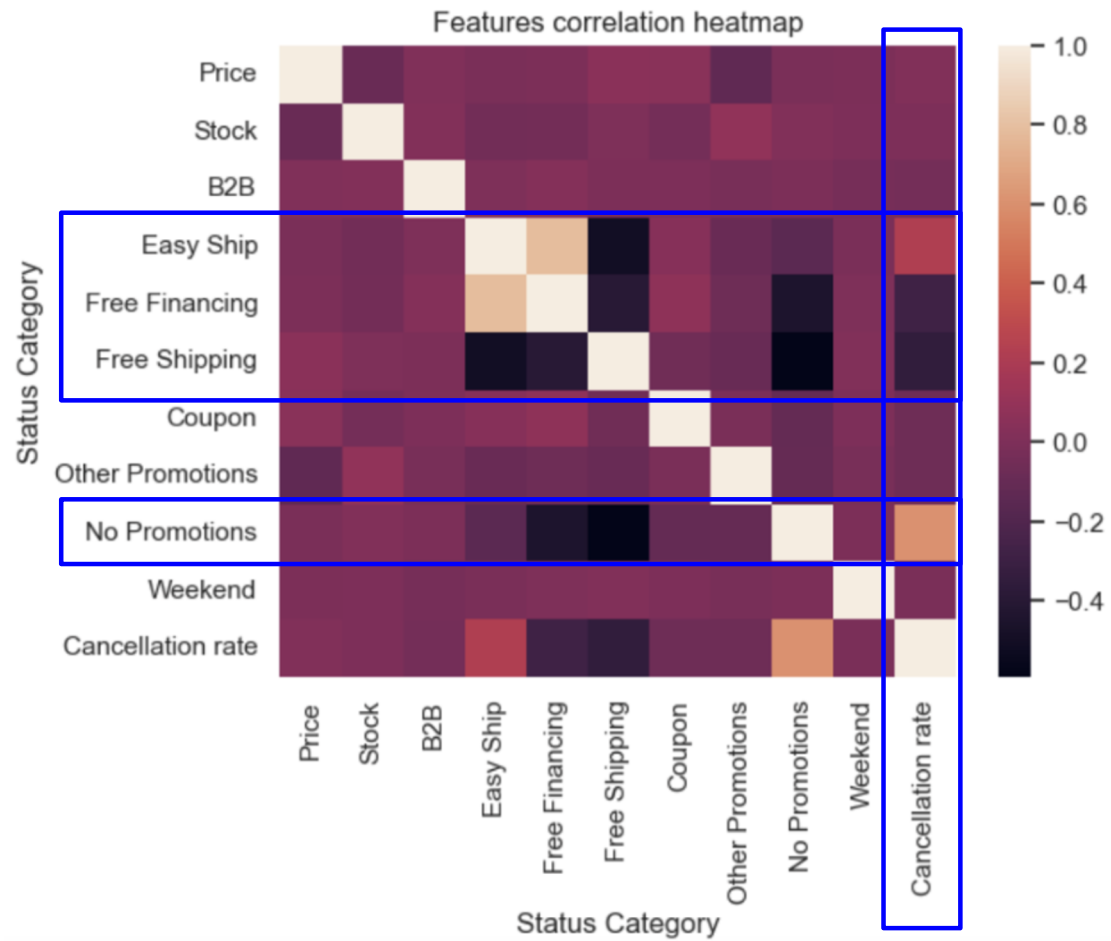


Cancellation Rate Per Product Category And Size



**Does size
matter?**

Correlation With Promotions



Cancellation Rate And Price



Modelling

Binary classification: 1 for cancelled order, 0 for shipped.

Imbalanced data: 14% of orders are tagged as Cancelled.

Recall as a main metric for performance assessment.

Pre-processing:

1. Label encoding.
2. Target-encoding for SKU.
3. One-hot encoding for other categorical features.
4. Scaling with RobustScaler.
5. Splitting into train (70%) and test(30%) data.
6. Resampling.

Selecting a model:

- Logistic regression
- KNN Classifier
- Random Forest
- Decision Tree

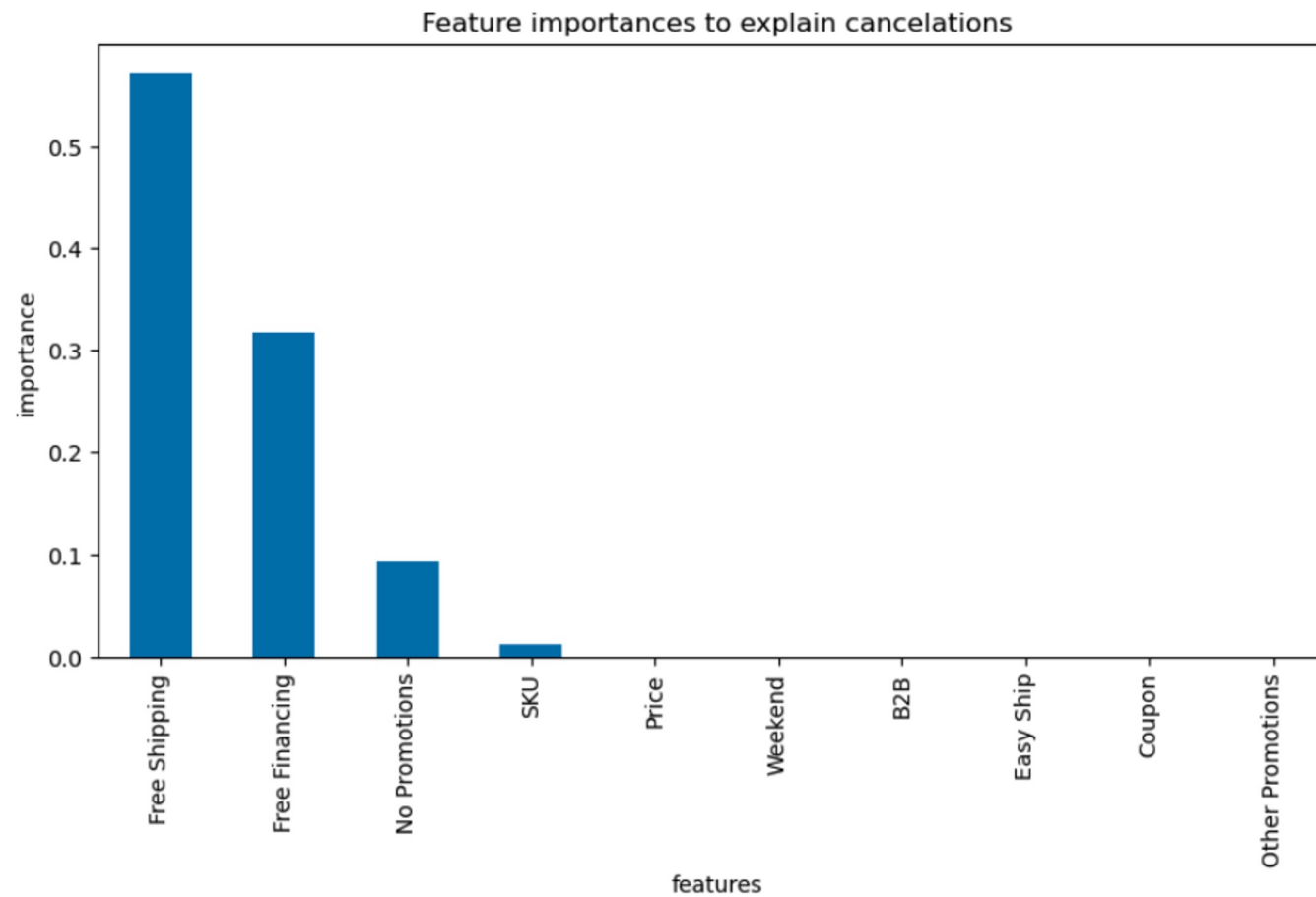
Hyperparameter tuning:

- Bayesian optimization.
- Using StratifiedKFold for cross validation (5 folds).

Models Comparison

Model	accuracy	precision <i>(for label 1)</i>	recall <i>(for label 1)</i>	f1-score <i>(for label 1)</i>
Logistic regression	0.81	0.42	0.94	0.58
KNN Classifier	0.8	0.4	0.84	0.55
Decision Tree	0.76	0.36	0.98	0.53
Random Forest	0.91	0.77	0.55	0.64

Final Results



Recall: 0.98
Precision: 0.31
F1 score: 0.47

Ideas To Improve The Model

1. Incorporate wait time



1. Include customer characteristics:

- Age
- Gender
- First-time customer vs repeat customer.
- Time spent on the website before placing an order.

