E-Commerce Shipping

June 2021

Winston Zhong

Business Problem

An e-commerce company wants to improve their on-time performance for their products' shipping. We will be looking at factors such as mode of shipment, weight, and product importance to see what the company can do to improve their product shipping performance. Shipping products on time is very important to the success of an e-commerce company so this is a crucial aspect of the company to improve. Findings from this will be shared with the leader of the logistics department to show what changes can be made to improve on-time performance.

We will also find out how the company can improve customer ratings. There are features about the product and customer that might have an effect on the customer rating. Having good customer ratings is another important feature for the company. Good customer ratings give the company a good image and more returning customers. There can be varying factors that affect customer ratings, so findings from this can be shared to leaders in the logistics, product development, and pricing departments.

Datasets

The dataset which contains 10999 entries will be downloaded from Kaggle. There are 12 columns in the dataset which range from information about the shipment, product, customer, and whether or not the order arrived on time.

Anticipated Approach

1. Wrangle data, make sure data is in the format we need and create new columns if necessary
2. EDA, look at our data distributions and do some hypothesis testing
3. Preprocess data, split data into training and testing sets, scale our data to prepare for modeling
4. Modeling data, use logistic regression and random forest methods to create classification models for whether orders were shipped on time. Use linear regression and random forest methods to create prediction models for customer ratings.

5. Use best models to predict what changes the company can make to improve on-time shipping performance and customer ratings.

Deliverables

1. All code for data wrangling, EDA, data preprocessing, and modeling
2. Final written report
3. Presentation slide deck