

Predicting Reordering for InstaCart Market

Xiao Zhang



InstaCart: online grocery ordering and delivery app



3 Million orders

Massive data!



Reorder predicting

- Make InstaCart easy to use by having products likely to reorder convenient to find
- Help grocery stores to optimize stocks: have products stock up when in demand
- Recommend relevant products to users to increase sales

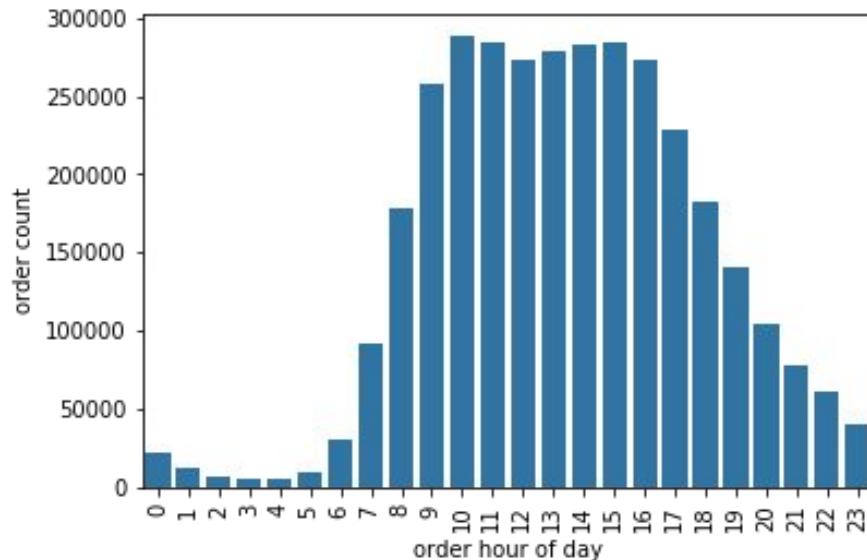
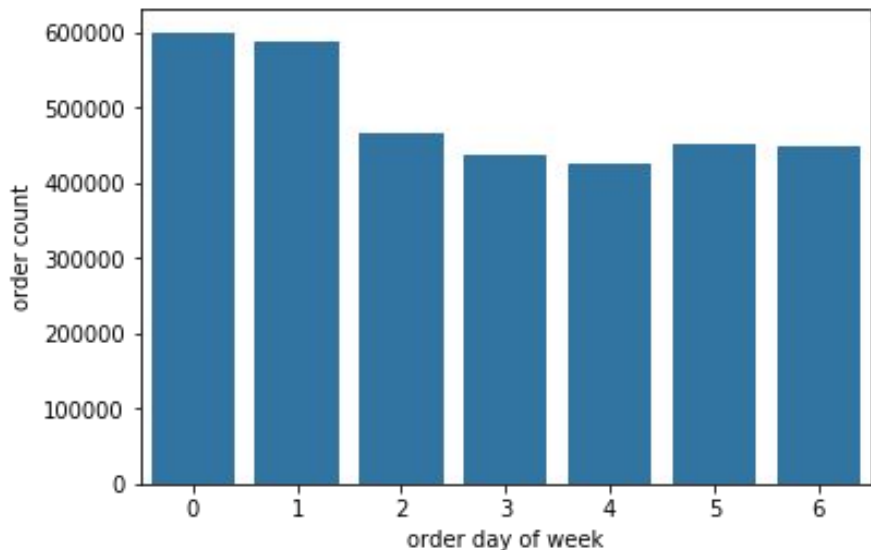


What can we learn from past orders?

- User properties
 - When the user orders, like what day of week, what time of day
 - How frequently
 - Does user prefer organic products
- Product properties
 - How often a product is ordered
 - Distribution of product order probability across week and day
 - Add to cart order
- User-Product properties
 - Frequency of a user purchasing a product

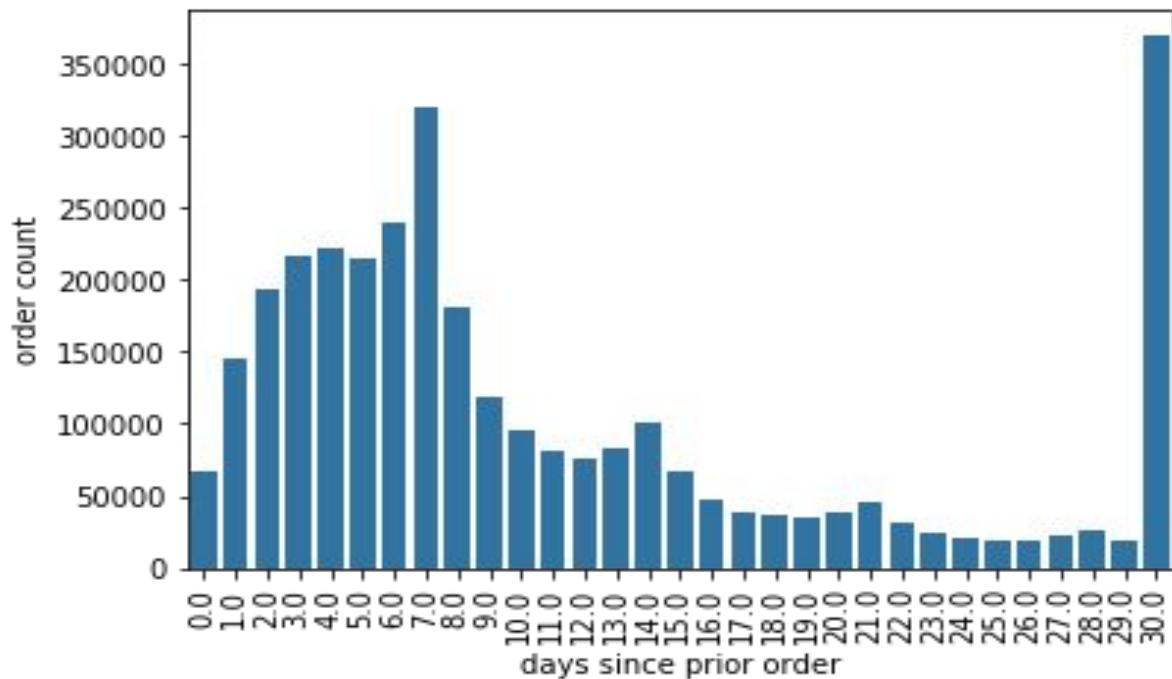


Orders mostly happen during the weekend, from 9am to 5pm



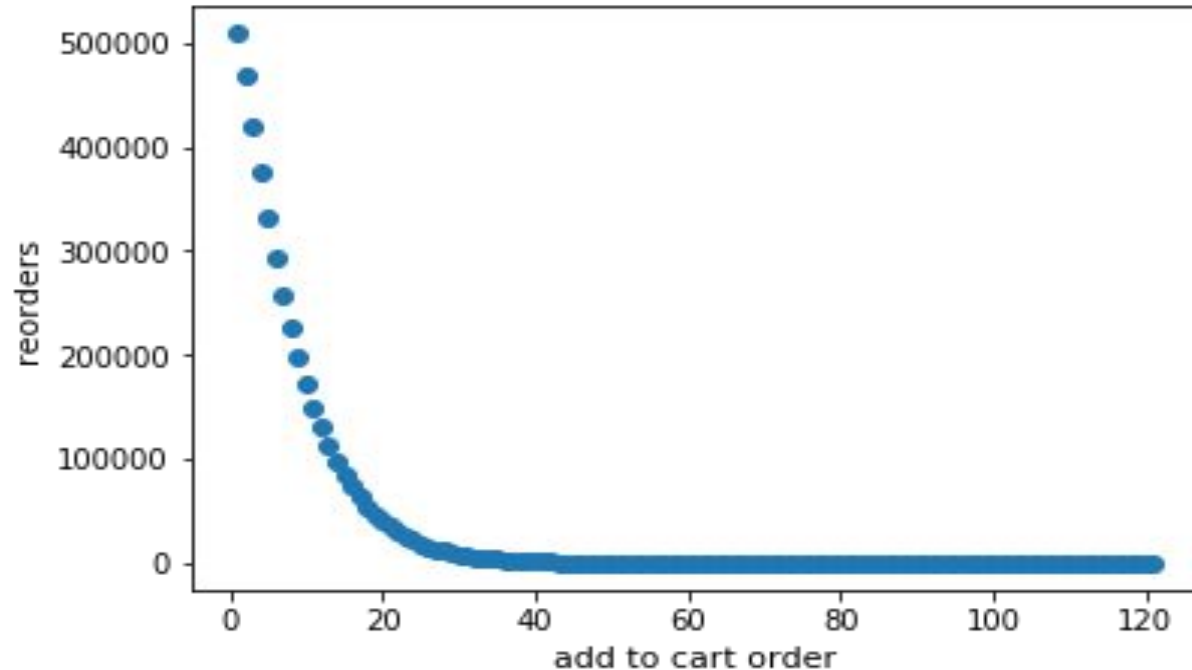


Users usually reorder after 7 days and 30 days



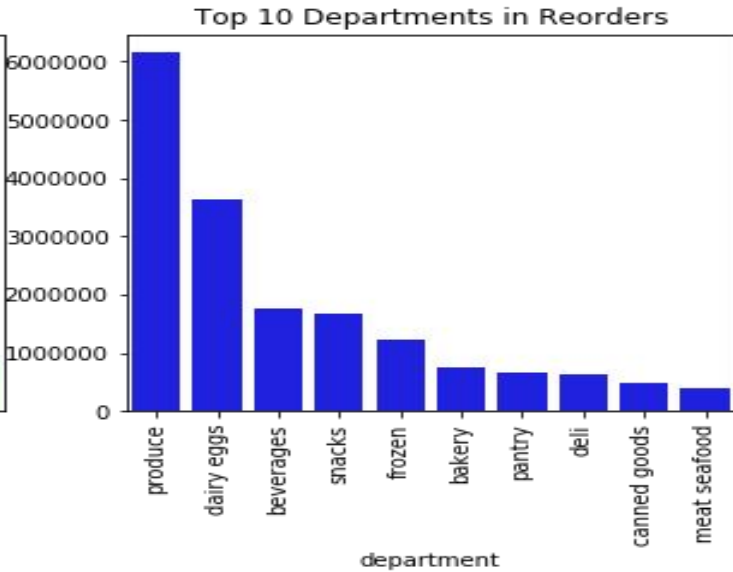
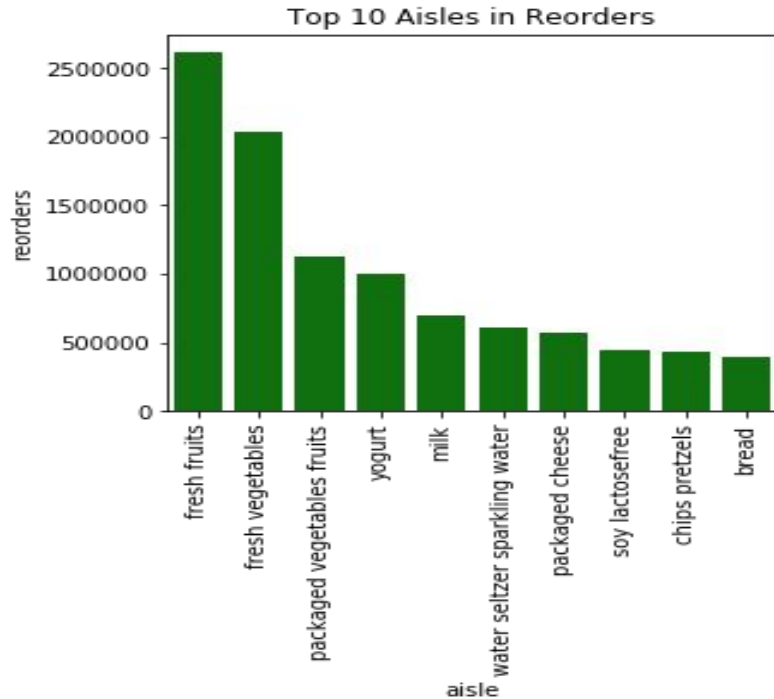


Customers tend to reorder the products they put in basket first

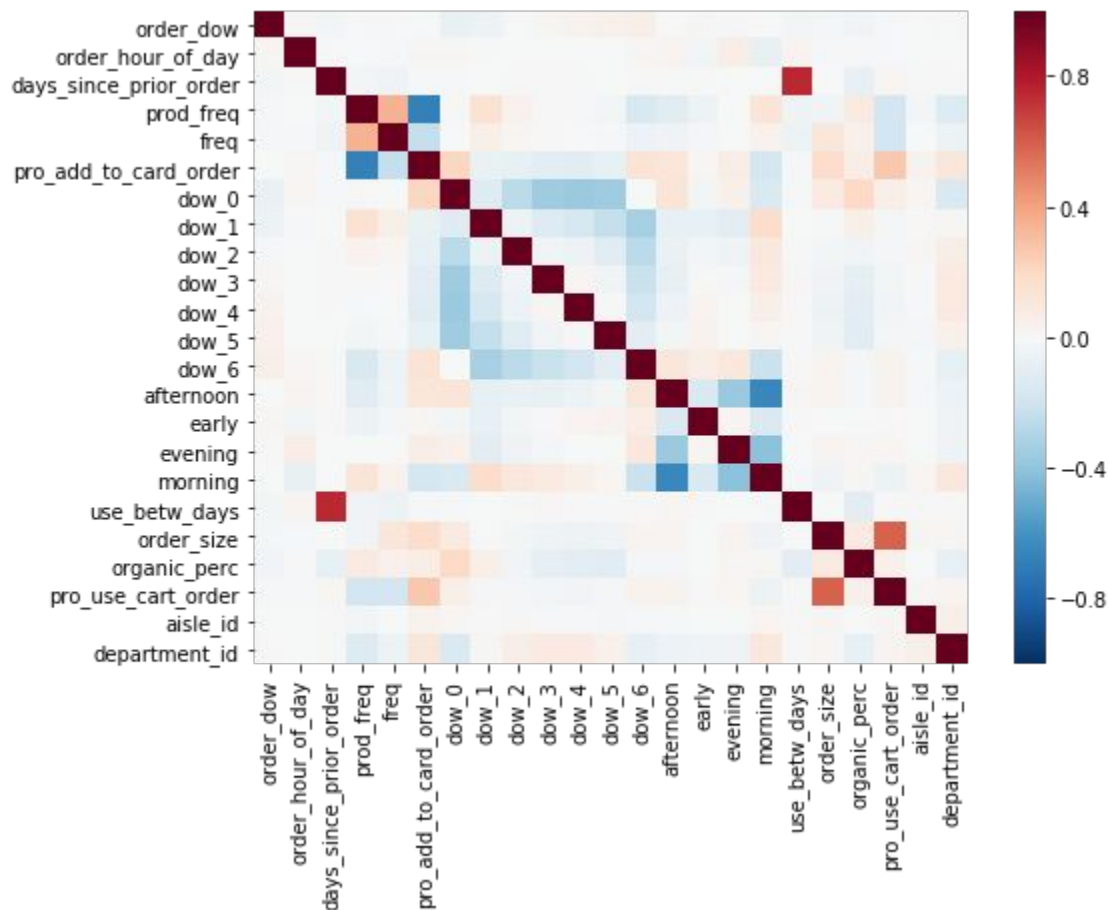


Most reordered aisles: fresh fruits and vegetables

Most reordered departments: produce and dairy eggs

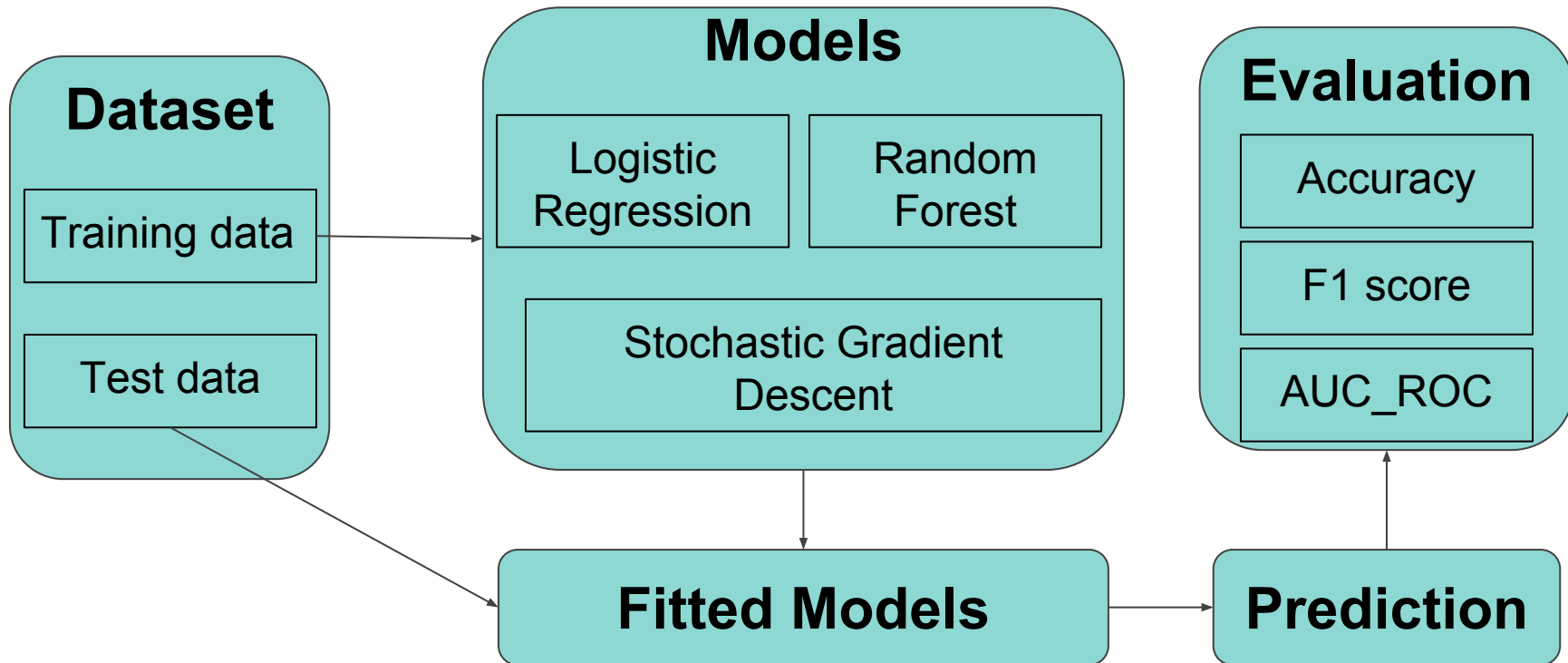


- 23 features representing four groups: user properties, product properties, user-product properties, date time information
- 2 features removed due to collinearity





Modeling





Results

- Logistic Regression and Random Forest Classification both perform well with training data (accuracy > 0.76 , $f1 > 0.78$, $auc_roc > 0.75$)
- Random Forest Classification has slightly higher $f1$ score than Logistic Regression
- $F1$ score on test data from both models are > 0.33 , compared to 0.4 of first rank Kaggle answer



Summary

- Both Logistic Regression and Random Forest Classification are able to predict reordering with decent performance
- Random Forest performs slightly better than Logistic Regression based on f1 score metric
- This study lays a good foundation of reordering prediction for InstaCart and provides helpful information to improve customer experience