Churn: Analysis & Solutions

Jess Kerlin
May, 2024

Agenda

- Preface
  - The Case Study
  - My Approach
- Business Presentation
  - Data Issues
  - Churn Predictor 1.0
  - Insights and Actions

The Case Study

The Data

- Data from over 5000 clients
- Churn Flag for each client
- 17 potential variables related to churn
- Columns are labeled with minimal descriptions

The Task

- Analyze the data
- Develop actionable insights and recommendations for Customer Experience (CX)

The Challenge

- Data Source Unknown
- To give useful insights and recommendations, you need to know the source, trust the data, know your context and audience

My Approach: Make a Story

- **Company Name:** E-Comm
- **Company Mission:** Be the #1 retail destination of the world

Questions with Answers

- What is this company's business model?
  - *Retail, we get margin on items we buy in bulk*
- Is there a subscription?
  - *No*
- How is churn defined?
  - *No E-Comm purchases for 1 month (April 2024)*
  - *No re-entry allowed (separate recovered category)*
- What does cashback mean in this context?
  - *We give discounts to clients on items we sell*
- How do we measure client value?
  - *Expected Lifetime Spend*
- Why are there nulls in the data?
  - *No one knows*
- Which stakeholders in the company are engaged?
  - *Product, CSM, Sales, Marketing … Anyone*

Variables with Vague/Unknown Units:

Tenure: *Months since first purchase*

CityTier: *Lower number allows higher markups*

WarehouseToHome: *Miles*

OrderAmountHikeFromlastYear: *No one knows*

CashbackAmount: *% Discount * 10*

Satisfaction Score: *Likert scale, Higher is better*

CC == Credit Card, COD == Cash On Delivery

All Mobile + Phone labels combined to Smartphone

Tackling Churn at E-Comm

E-Comm's Core KPI : Churn

The Problem

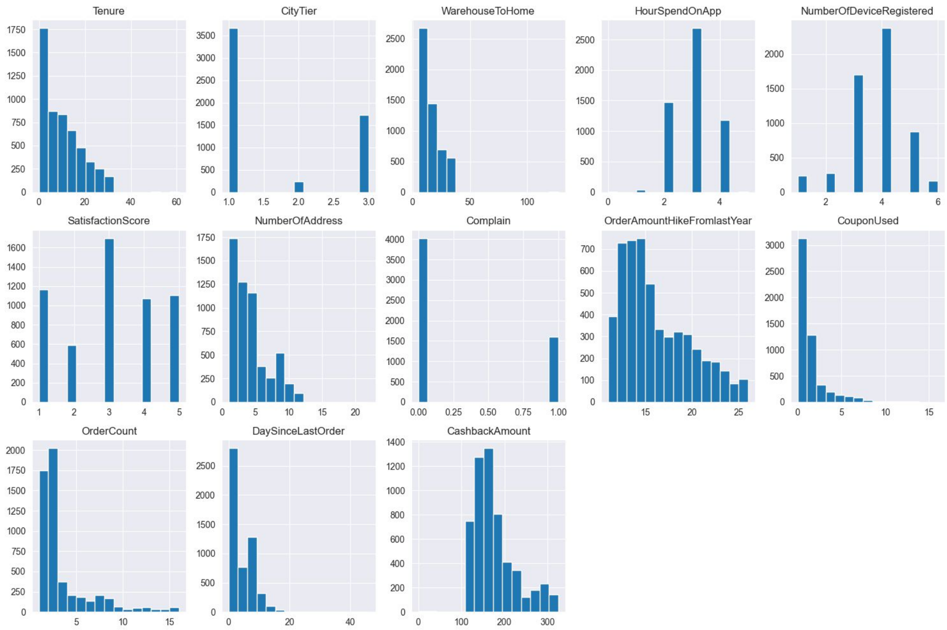
- Customer churn is the largest source of lost revenue at this company
- Reducing churn is the #1 priority for this company across the board this year
- **17%** of customers churn month-to-month

Our Goals Today

- Highlight data issues where action needs to be taken
- Provide Product and Customer Support with a tool to predict which clients will churn, so they can implement targeted interventions
- Provide insights into where to look to improve our customer experience

The Data

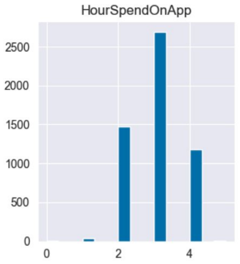
- Data from over 5000 clients that bought at least 1 item **every** month up to and including March 2024
- Did they buy in April (Retained) or not buy (Churned)?
- 17 potential variables related to churn
  - User Tenure
  - Preferred Login Device
  - City Tier
  - Distance from Warehouse to Home
  - Preferred Payment Method
  - Gender
  - Hours Spent on App
  - Number of Devices Registered
  - Preferred Order Category
  - Satisfaction Score
  - Marital Status
  - Number of Shipment Addresses
  - Was there a complaint in the last month?
  - Order Amount Increase for Last Year
  - Coupons Used
  - Number of Orders
  - Days since Last Order
  - % Cashback Discount on Items Bought



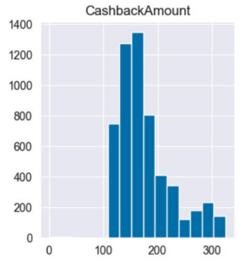
Data Issues

Strange Distributions

All clients spend 2-5 hours on the app?
Friction at sign-up?
Usage capped?



Are all of our discounts between 10 and 30%?
Strategic?

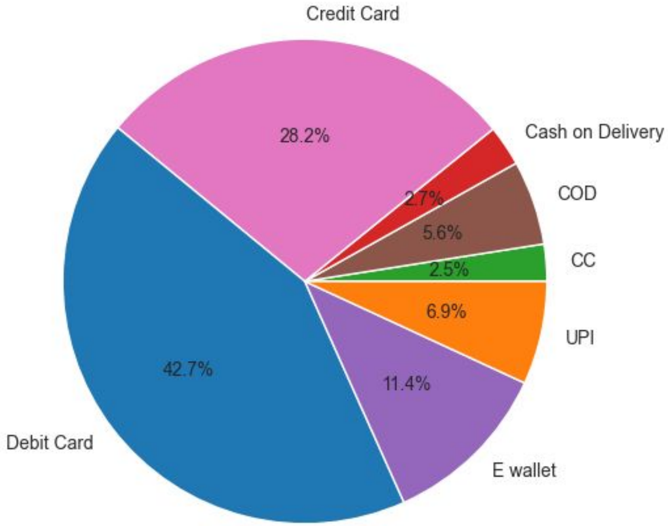


Nulls In The Dataset

	Column Name	Number of Nulls
0	CustomerID	0
1	Churn	0
2	Tenure	264
3	PreferredLoginDevice	0
4	CityTier	0
5	WarehouseToHome	251
6	PreferredPaymentMode	0
7	Gender	0
8	HourSpendOnApp	255
9	NumberOfDeviceRegistered	0
10	PreferedOrderCat	0
11	SatisfactionScore	0
12	MaritalStatus	0
13	NumberOfAddress	0
14	Complain	0
15	OrderAmountHikeFromlastYear	265
16	CouponUsed	256
17	OrderCount	258
18	DaySinceLastOrder	307
19	CashbackAmount	0

Duplicate Labels

Distribution of Preferred Payment Mode across clients with Tenure of 2+ months

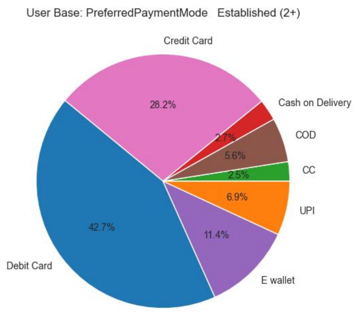


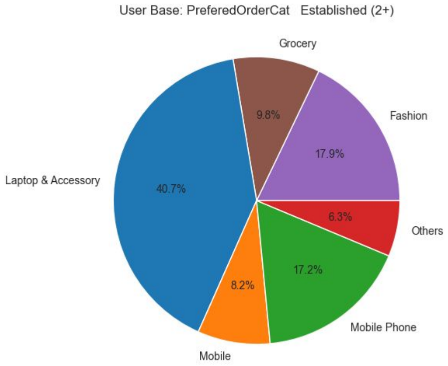
Data Issues: It goes deep

A complex data problem

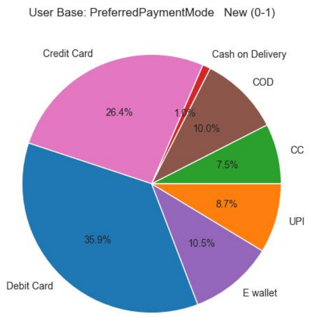
- Tenure “null” users are more likely to Prefer to pay with ‘CC’ than ‘Credit Card’
- Tenure “null” users are overwhelmingly more likely to order ‘Mobile’ or ‘Mobile Phone’
- Is our client base changing over time?
- Labeling is unified for the analysis, but needs to be investigated in the product

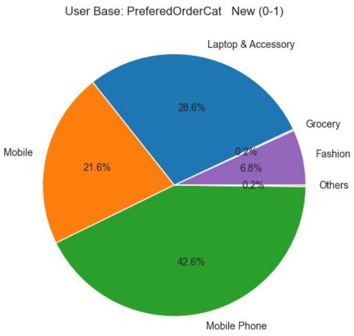
Tenure 2+



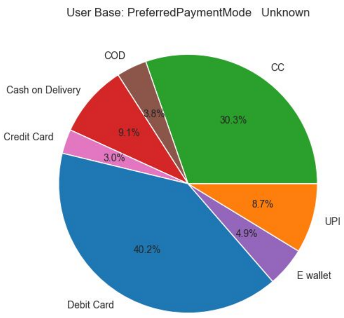


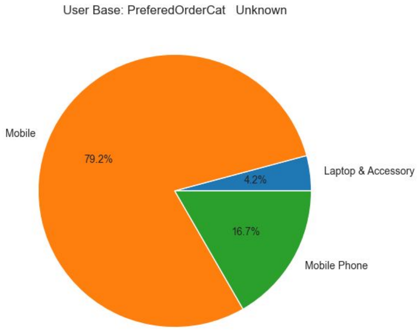
Tenure 1-0





Nulls





Introducing Churn Predictor 1.0

The Data

- Data from over 5000 clients
- Churn Flag for each client
- 17 potential variables related to churn
    - User Tenure
    - Preferred Login Device
    - City Tier
    - Distance from Warehouse to Home
    - Preferred Payment Method
    - Gender
    - Hours Spent on App
    - Number of Devices Registered
    - Preferred Order Category
    - Satisfaction Score
    - Marital Status
    - Number of Shipment Addresses
    - Was there a complaint in the last month?
    - Order Amount Increase for Last Year
    - Coupons Used
    - Number of Orders
    - Days since Last Order
    - % Cashback Discount on Items Bought

The Model

Random Forest Classification Model

The Result

97% of Churners Captured

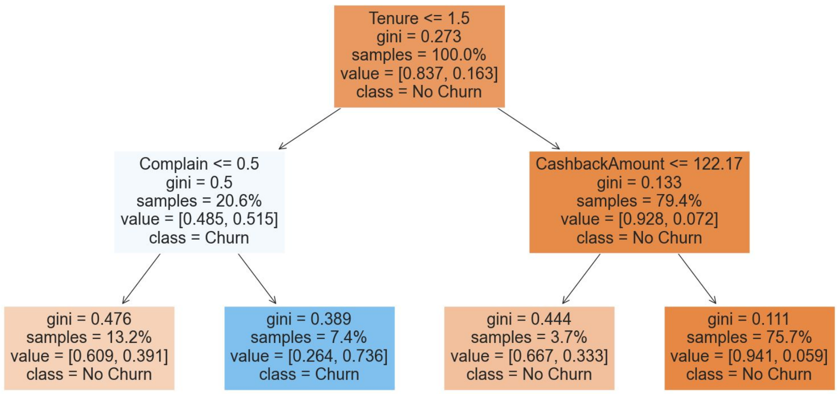
3% False Alarm Rate

86% of those labeled as Churn will Churn

	precision	recall	f1-score	support
0	0.99	0.97	0.98	933
1	0.86	0.97	0.91	193
accuracy			0.97	1126
macro avg	0.93	0.97	0.94	1126
weighted avg	0.97	0.97	0.97	1126

What is Random Forest Classification?

Single Decision Tree



Random Forest Classification*

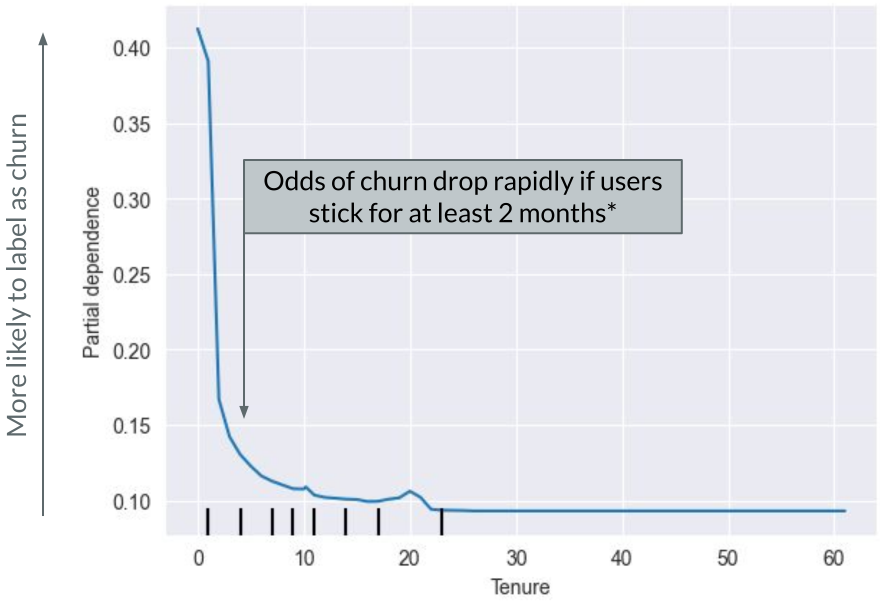
*Details:
100 estimators, No max depth, class weight balanced
Threshold set to 0.3 to weight toward Churn recall
Experiments find maximal performance
All Features used except customerID, nulls impute mean

Random Forest: Bonus Info

- Random Forest Classification offers additional insight into what matters:
  - Feature Importance
    - Which feature has the biggest influence on whether a user will be predicted to churn?
  - Partial Dependence
    - What direction is the feature typically pushing?

Feature: Tenure
Feature Importance: 0.18 (highest score, Rank 1)

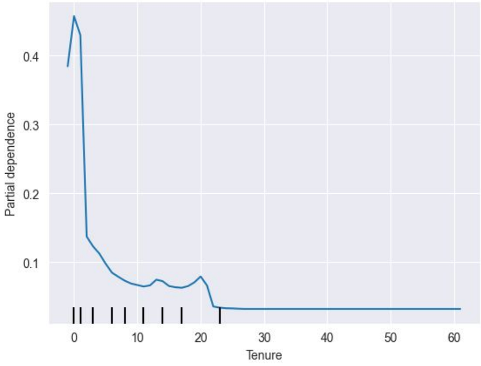
Partial Dependence Plot



The Top 3 Predictors

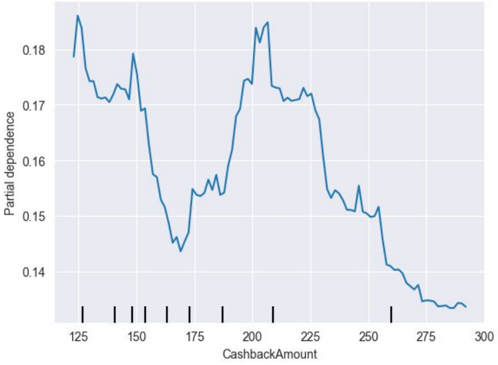
1 - Tenure

- Long-term customers stay loyal*



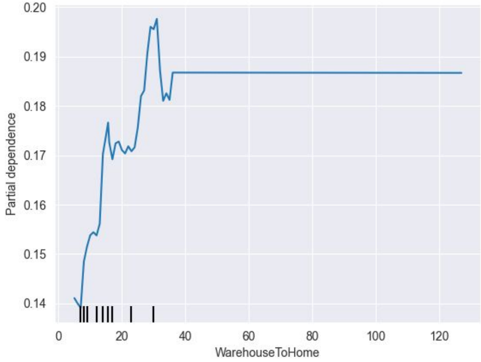
2 - Cashback Amount

- Cashback is wavy*



3 - Warehouse Distance

- Long distance from warehouse is bad*

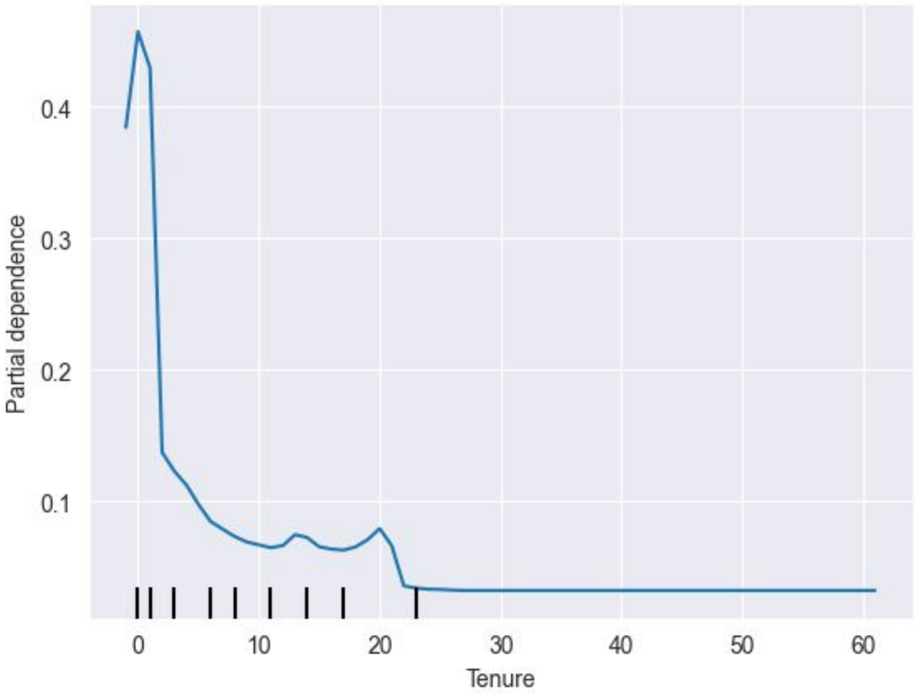


*Caveats

Why does Tenure matter?

A few considerations

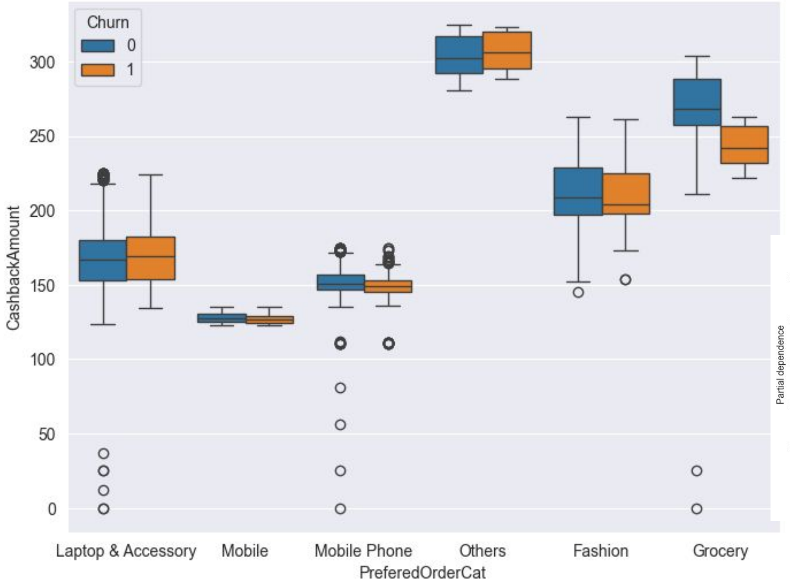
- Customer loyalty
- Late item delivery?
- Our latest users are different?
- No re-entry of churned customers in our definition!
  - Where do they go?
  - How do they fit?
- Need even more data to disambiguate:
  - Past snapshots of Churn
  - Better data labeling
  - More product funnel info
  - Complaint analysis
  - Delivery analysis

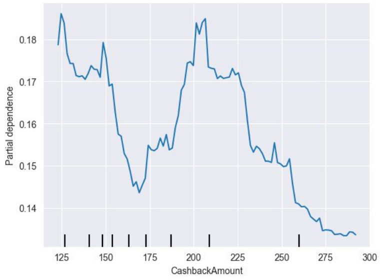


Why is cashback so complicated?

The product and users are complex

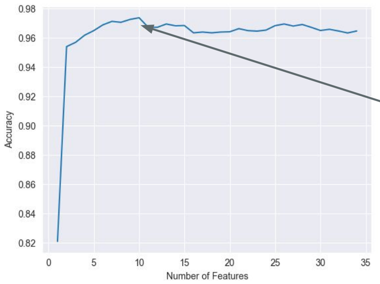
- Cashback is defined by our product (different ranges offered in different product categories
- Cashback is determined by our users (whether or not they choose to purchase particular items on sale)
- Need more data to disambiguate ...
  - Purchase pricing
  - Itemized deals
  - Better labeling
  - Cohort analysis from sale periods
- Are all clients equally valuable? Could drive further modeling of cost/benefit of churn reduction





Insights: Feature Importance Top 10

Rank	Feature	Impact Directionality (from Partial Dependence)
1	Tenure	Longer is better
2	Cashback Amount	Complex
3	Warehouse Distance	Lower is better
4	Days Since Last Order	~3 to 10 is best
5	**Did User Complain?**	No is better
6	**Number of Addresses**	Fewer is better
7	Order Hike From Last Year	~14 to 21 is best
8	**Satisfaction Score**	Lower is better
9	**Order Count**	Lower is better
10	**Number of Registered Devices**	Lower is better



Accuracy stops improving after ~10 features

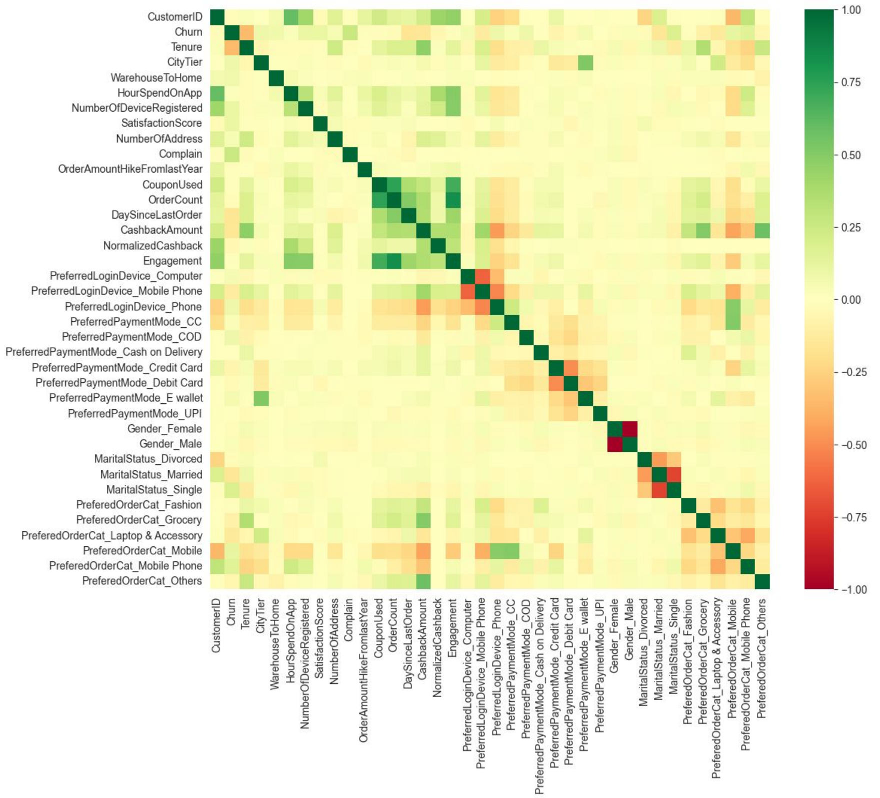
Follow-ups

- Number of Addresses / Order Count
  - Friction in the interface?
  - Short-term deal customers getting gifts for holidays?
- Satisfaction Score
  - When in the funnel do we measure this?
  - Reconsider the measure?
- Number of Registered Devices
  - Poor multi device integration in the product?
  - Error tracking clients in churn analysis?
- Complaints
  - What are these complaints?

The Rest of the Features

Many correlations, few relevant

- While there are many interdependencies between features like purchase mode, purchase category and marital status, they not in themselves predictive of churn
- To be revisited after feedback from BI/Product



Next Steps

- Cashback and warehouse distance are key areas to investigate client experience
- Product, CS and BI to follow up of data issues highlighted here, especially related to labeling and Tenure
- DS to provide current client churn risk labels to Product, Marketing and CS as authorized and needed
    - Re-engagement campaigns
    - Calls for feedback
    - Better satisfaction survey

Thanks for your attention

Case Study Appendix

<u>Customer Churn Data</u>:

CustomerID: Unique customer ID

Churn: Churn Flag

Tenure: Tenure of customer in organization

PreferredLoginDevice: Preferred login device of customer

CityTier: City tier

WarehouseToHome: Distance in between warehouse to home of customer

PreferredPaymentMode: Preferred payment method of customer

Gender: Gender of customer

HourSpendOnApp: Number of hours spend on mobile application or website

NumberOfDeviceRegistered: Total number of deceives is registered on particular customer

PreferedOrderCat: Preferred order category of customer in last month

SatisfactionScore: Satisfactory score of customer on service

MaritalStatus: Marital status of customer

NumberOfAddress: Total number of added added on particular customer

Complain: Any complaint has been raised in last month

OrderAmountHikeFromlastYear: Percentage increases in order from last year

CouponUsed: Total number of coupon has been used in last month

OrderCount: Total number of orders has been places in last month

DaySinceLastOrder: Day Since last order by customer

CashbackAmount: Average cashback in last month

Data Munging Details

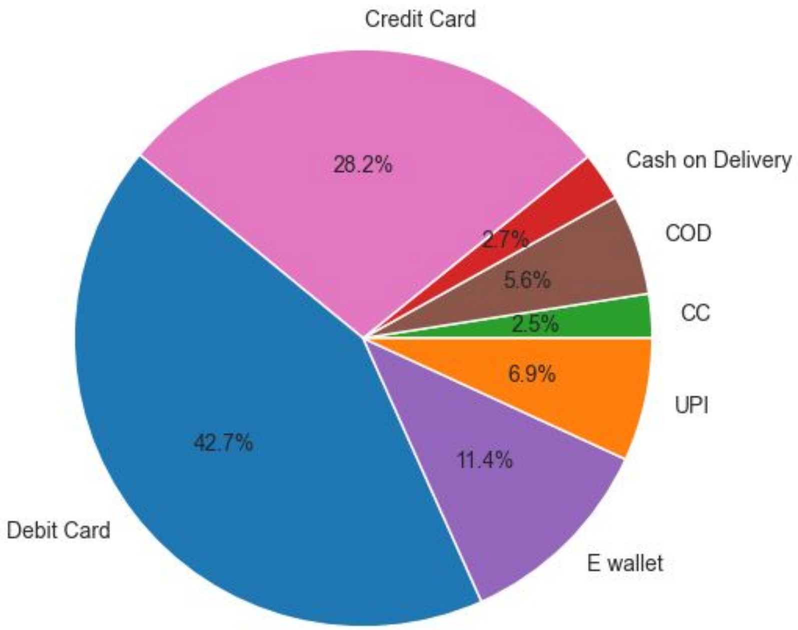
Spelling errors in columns corrected
All classes with the word Phone or Mobile changed to “SmartPhone”

CC → Credit Card
COD -> Cash on Delivery

Nulls set to -1 for RF:
Tenure
HourSpendOnApp
DaysSinceLastOrder
OrderCount

Nulls set to mean:
WarehouseToHome
OrderAmountHikeFromlastYear
CouponUsed

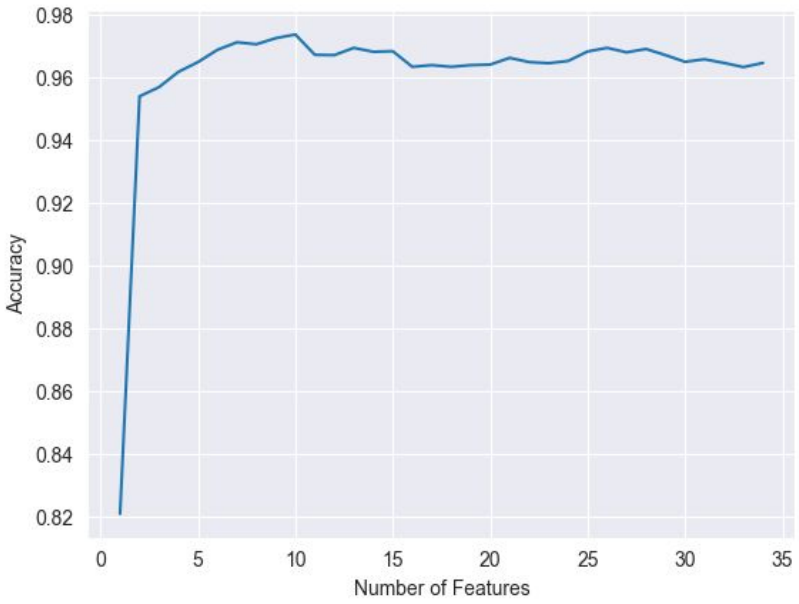
Note: These changes did not dramatically impact model performance or feature selection



Random Forest Details

*Details:
100 estimators, No max depth, class weight balanced
Threshold set to 0.3 to weight toward Churn recall
Experiments find maximal performance revealed accuracy peaks around 10 features

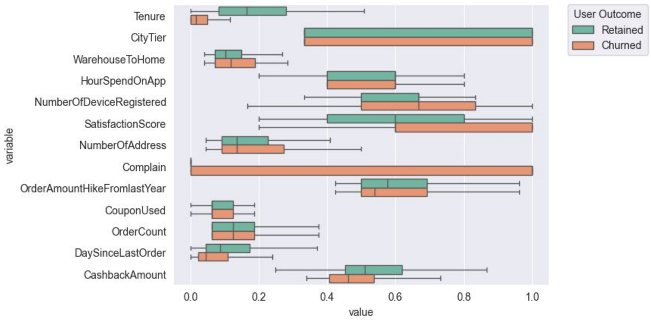
100 estimators, No max depth, class weight balanced
Threshold set to 0.3 to weight toward Churn recall, since churn costs much more than intervention.



Why partial dependence?

The problem is complex

- Some useful features are binary or had non-linear relationships
- Single decision trees and linear models failed to accurately capture churn accurately



Example: Fashion