

Making Spam Meat Again

Leveraging Targeted Email Marketing to Reduce Churn

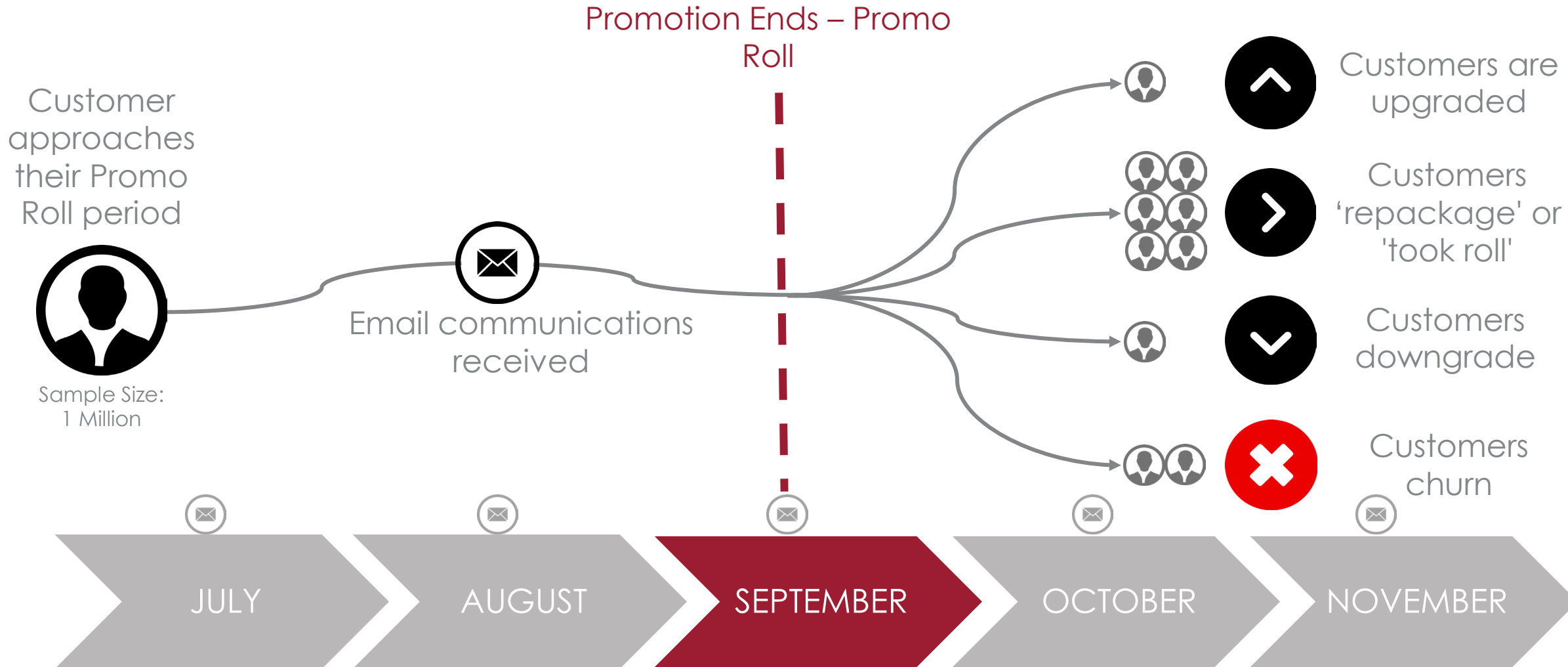
xfinity

MIT
MANAGEMENT
SLOAN SCHOOL

AGENDA

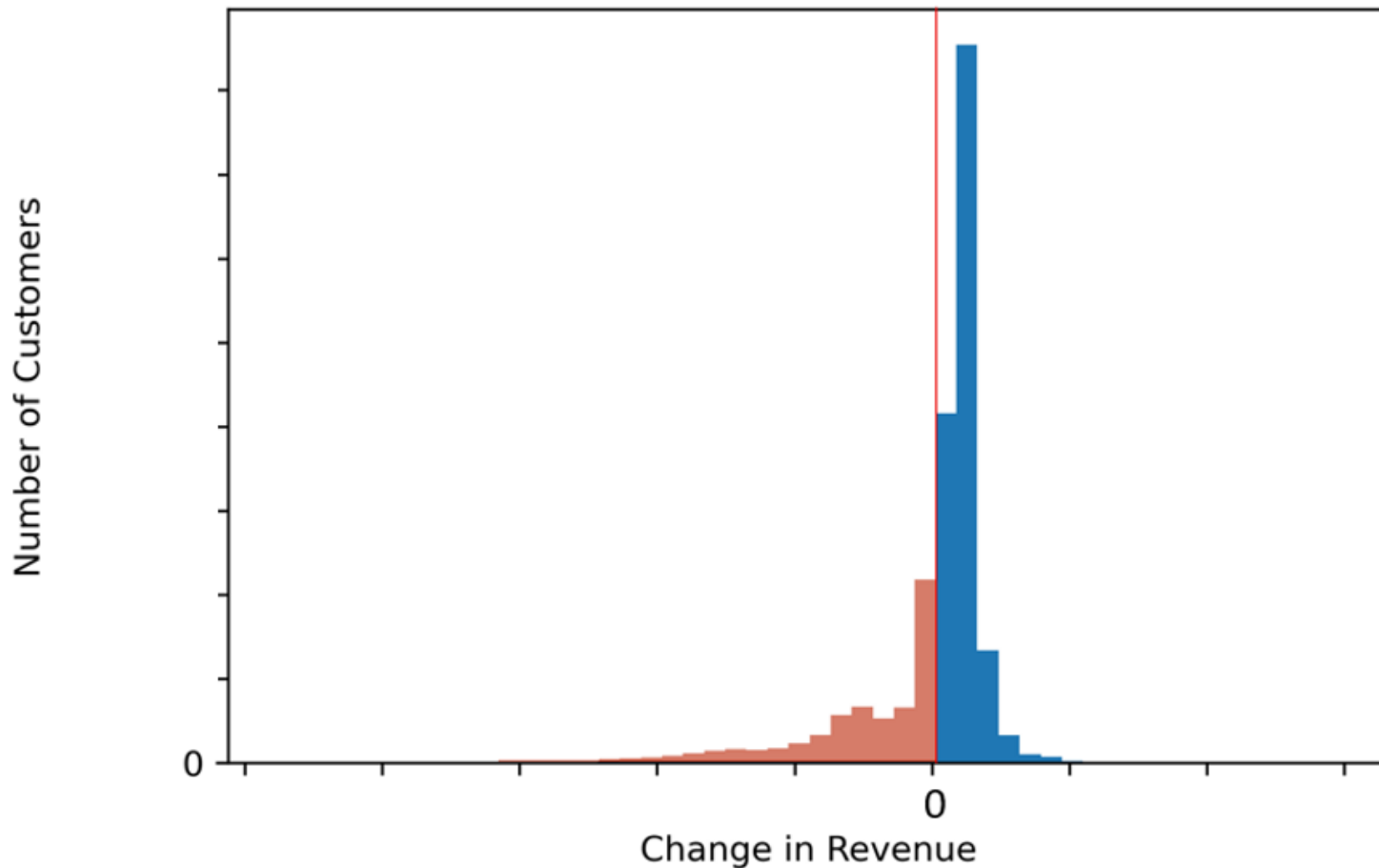
1. Project Overview
 - A. Core Goals
2. Emails and Churn
 - A. Promo Roll Churn rate analysis and Segmentation
 - B. Email Churn Rate Prediction
3. Prescription and Results
 - A. Prescriptive Optimization Model
 - B. Deliverable Handover and Transition

1.1 DATA OVERVIEW



1.2 WHY PROMO ROLL?

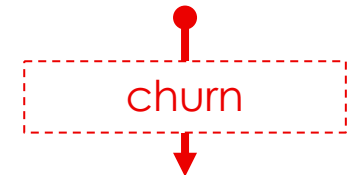
Actual Monthly Recurring Charge Change



Average expected MRC increase/customer:

21%

Assuming no churn



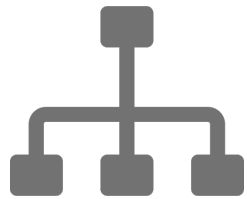
Actual Average MRC increase per customer post-churn:

2%

1.3 OUR GOALS



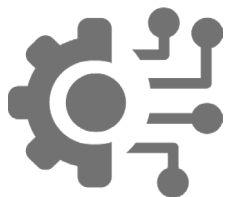
1: Investigate Churn likelihood and identify non-Email drivers behind it



2: Use clustering to identify **segments** within Promo Roll customers

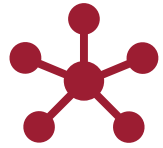


3: Investigate the impact of Emails on Churn, and develop a model to **predict Churn from Emails**



4: Develop a **prescription framework** to recommend email mix strategy to each segment to minimize Churn and maximize Post Roll MRC

2.1 PROMO ROLL SEGMENTATION



K-means based on 12 selected features – k=6



Decision tree for ease of business use



Basis for optimization – personalize email mixes



Clusters centroids as “sample participants”

Cluster	Churn Δ
1	+13.7%
2	-0.9%
3	-2.2%
4	-2.4%
5	-4.1%
6	-6.3%

2.2 EMAIL CHURN RATE PREDICTION

Predict churn using email categories and counts

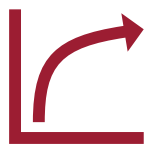


1.6 million customers



Data Includes:

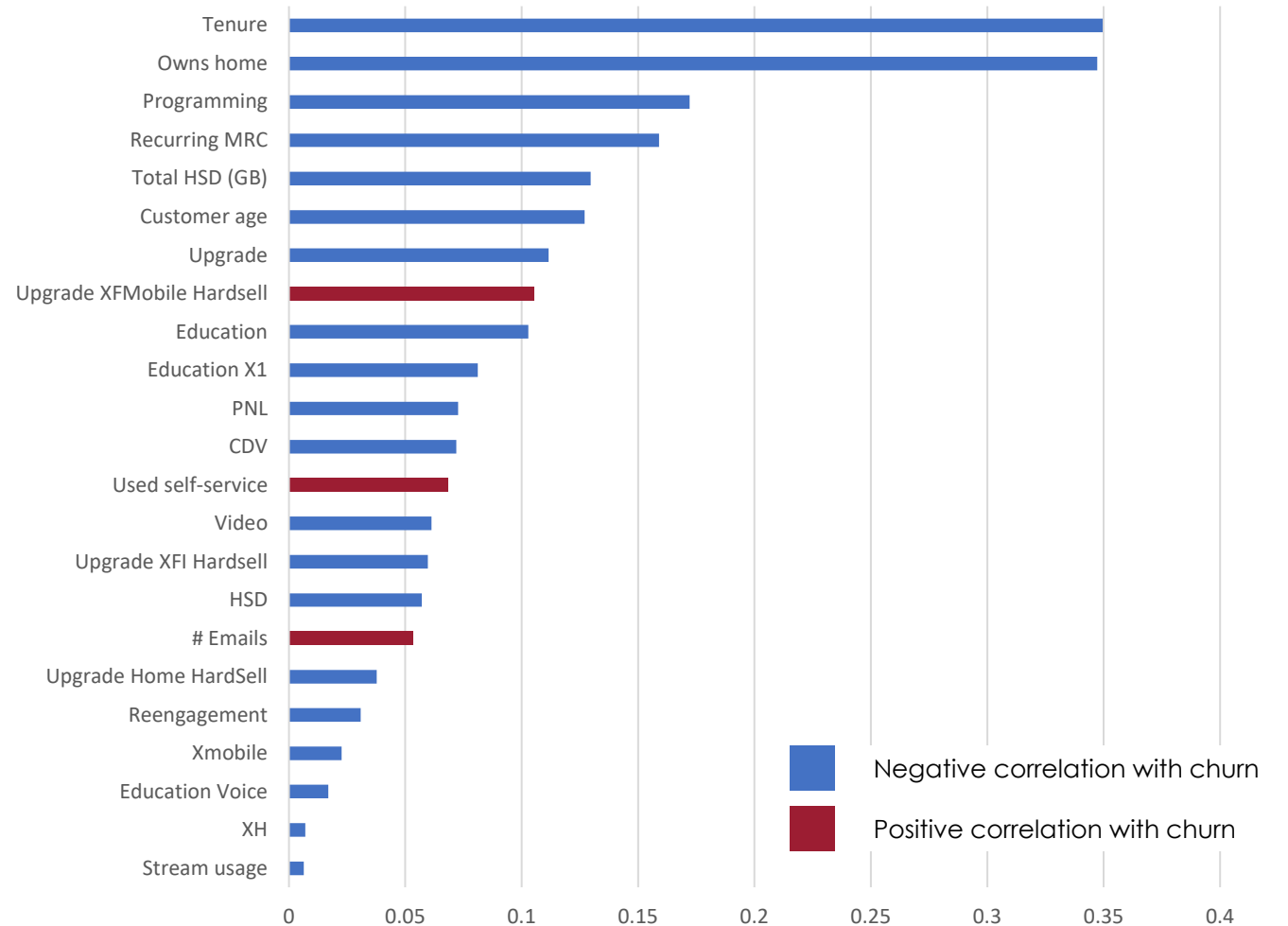
- # of total emails
- Emails by key programs
- Selected demographic features



Final model: Logistic Regression

- AUC: 0.726
- Accuracy: 0.893

Importance



Blue: Negative correlation with churn
Red: Positive correlation with churn

3.1 EMAIL MIX OPTIMIZATION

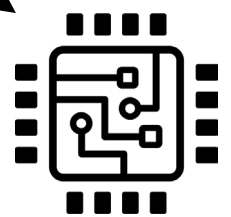
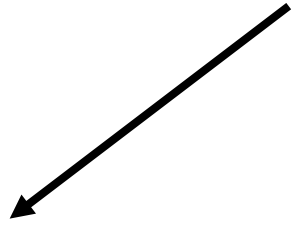
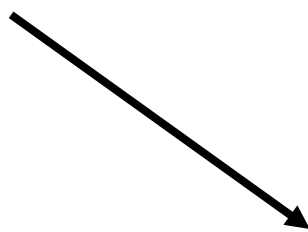
Segment Assignment



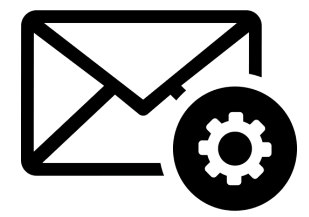
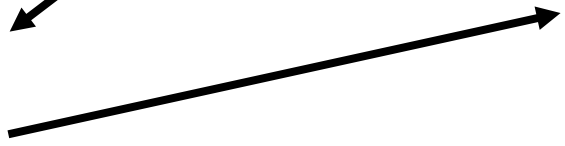
Produce and Usage Data



Predictive Model



Cluster Centroid Optimization Formulation



Email Mix Strategies with assignments for each rolling cluster

3.2 OPTIMIZATION FORMULATION

Objective Values:

$$\text{Min}_x \sum_c (\beta_{c,c} x_{c,c} + \gamma_{c,c} z_{c,c} + \rho_c \cdot (s_c + \sum_j x_{c,j})) + \sum_t \alpha_t x_{t,t}$$

[Log Odds] [Timing]

Key Variables/Inputs:

- x_{cjt} - sending an email j to cluster c in week t
- z_{ij} - known centroid demographic value j for cluster c
- β_{cj} - coefficient for cluster c for email campaign j
- γ_{ck} - coefficient for cluster c for demographic k
- α_t - coefficient for emails in week t

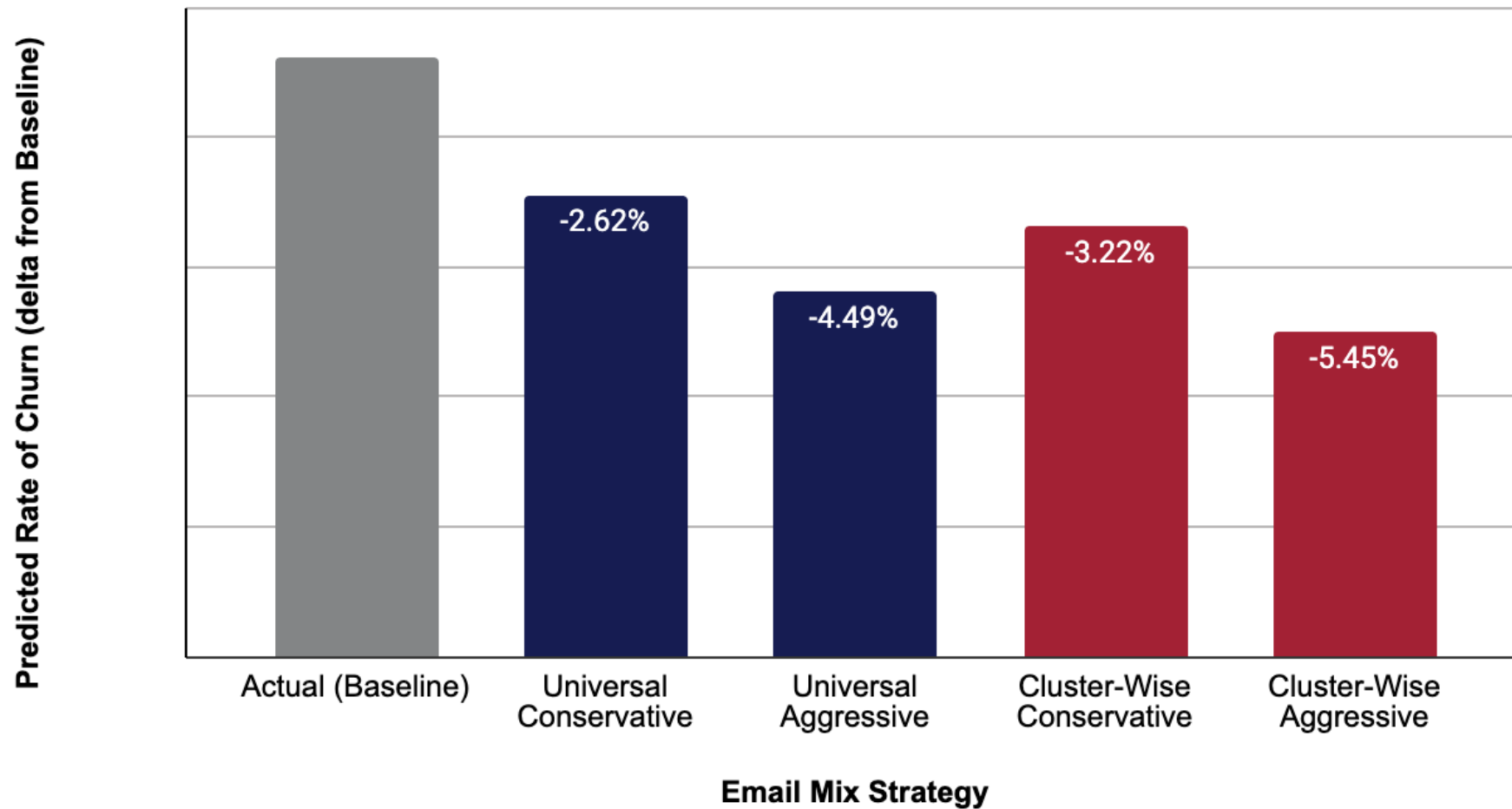
Constraints:

- Restrict the total number of emails that can be sent in any one week
- Restrict the total number of emails sent in any one category
- Restrict the total number of emails from one category sent in any two-week period

Exact constraint values malleable for different strategies.

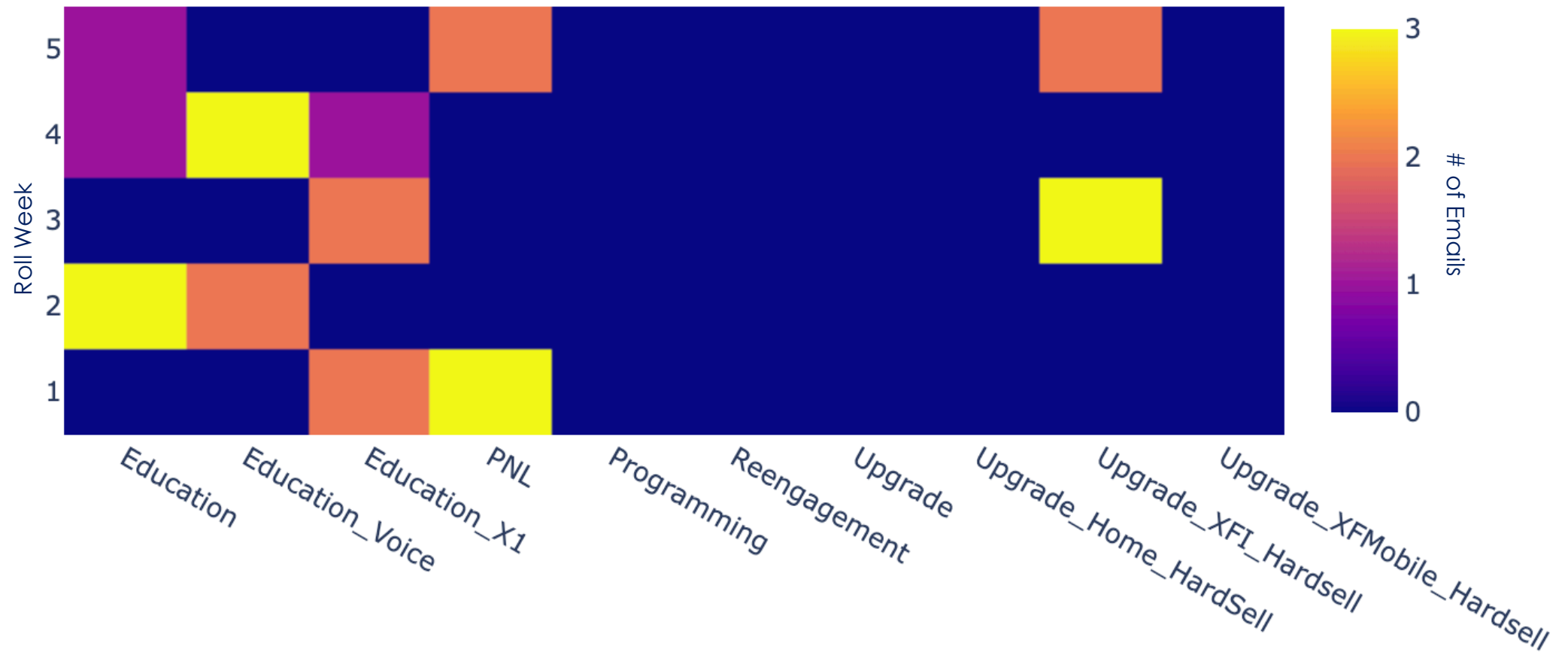
3.3 EMAIL MIX OPTIMIZATION

Test Set Optimization Scenario Results



3.4 EXAMPLE EMAIL MIX

Aggressive Cluster 1 Email Assignments



1
2
Prescription

3.5 BUSINESS IMPACTS



Uncovered differentiation between groups of Promo Roll customers through segmentation



Analysed the role of emails in impacting churn likelihood and outlined experiments for *continued inference* and *impact measurement*



Developed a Prescriptive Framework that provides email mix strategies that reduced churn by over **3 percentage points** in offline testing

3.6 DELIVERABLE HANDOVER

Concise Handover Plan:



Technical Stakeholder: EBI Team

- Regular update meetings and input as the project has progressed
- Deliverable code documented for handover
- Email mix strategies and methodology employed
- Testing framework for future development



Business Stakeholder: LCM Team

- Consulted with at project inception
- Final results meeting to convey key insights
- Recommendations of experimentation on random emailing for causal inference

xfinity

MIT
MANAGEMENT
SLOAN SCHOOL

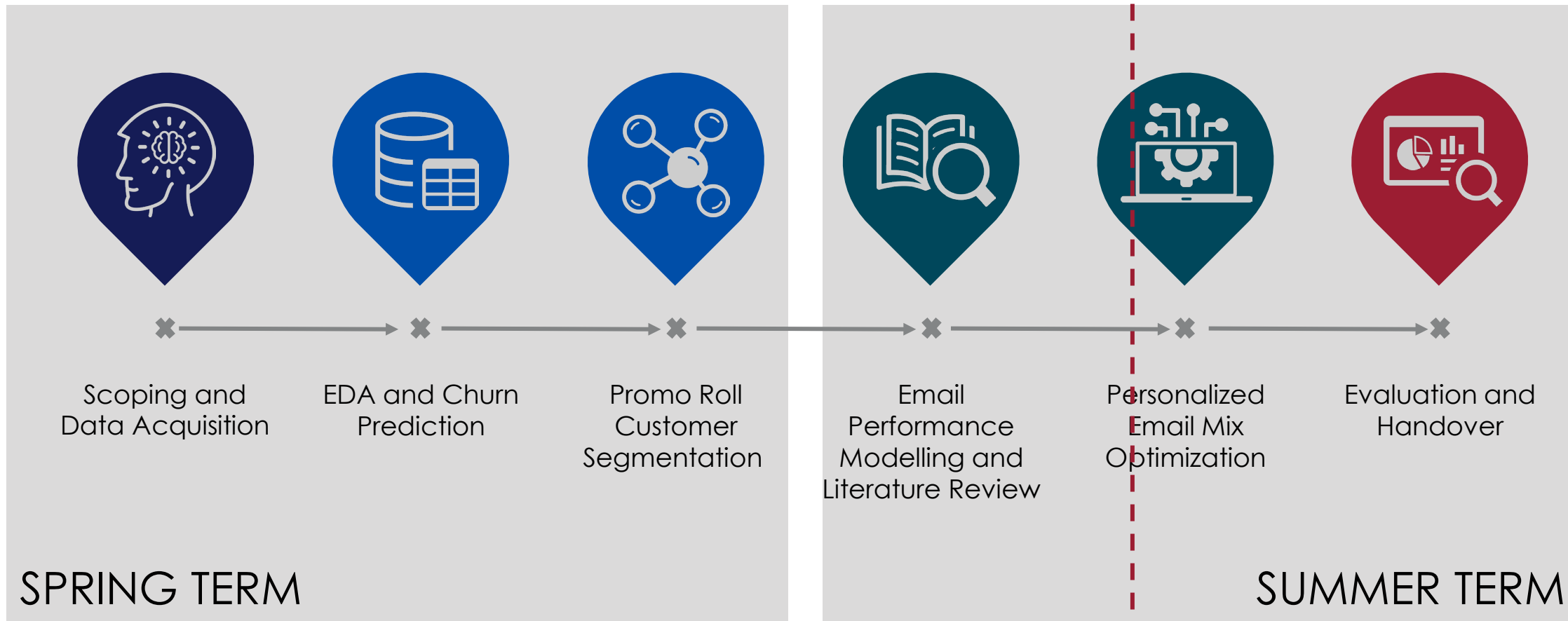
APPENDIX

APPENDIX 0: BRIEF ABBREVIATIONS DICTIONARY



- EBI Team: Enterprise Business Intelligence, our key technical and executive stakeholders at Comcast
- LCM Team: LifeCycle Marketing team, our primary business stakeholders at Comcast
- XF: Xfinity, product/marketing line for certain Comcast products
- EDA: Exploratory data analysis
- MRC: Monthly recurring charge, amount each customer is billed each month
- RGU: Revenue Generating Unit

APPENDIX 1.1 SIMPLIFIED PROJECT TIMELINE



More detailed task breakdown in appendix

CURRENT PROGRESS:

APPENDIX 1.2: DETAILED PROGRESS BREAKDOWN



Today

Deliverable	Task	Part-Time				Break	Full-Time				End	Progress											
		February		March		April		May		June			July		August								
		W1	W2	W3	W4	W1	W2	W3	W4	W1	W2		W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	
Initial Work	Project Scoping	■	■	■	■																●	100%	
	Stakeholder Meetings		■	■	■	■																●	100%
	Data Acquisition			■	■	■	■															●	100%
	EDA				■	■	■	■	■													●	100%
Churn Prediction Modelling	Model Development					■	■	■	■	■												●	100%
	Model Evaluation and Assessment						■	■	■	■	■											●	100%
Promoroll Segmentation through Clustering	Clustering of PromoRoll base							■	■	■	■											●	100%
	Cluster-Based Churn Prediction								■	■	■	■										●	100%
	Clustering-Segmentation Comparison									■	■	■	■									●	100%
Email Churn Prediction Modelling	Literature Review								■	■	■	■	■									●	100%
	Email Churn Prediction										■	■	■	■	■							●	100%
Prescriptive Modelling	Optimization model for Prescription											■	■	■	■	■						●	100%
	Evaluation/Testing of Optimization Model													■	■	■	■					●	100%

APPENDIX 2.1: DETAILED CHURN PREDICTION DATA



- Promo Roll was reduced to a binary (True-False) modelling problem around Churn Rate
- Numerous models (xgboost, cart, logistic regression, among others) were trialled
- Challenges involving unbalanced dataset size and signal strength

Modelled Variables:

Demographic Features

Age, Tenure with Comcast
Home Ownership Status (rent vs. own, etc.)
Inferred Family Information
...

Product Features

Product Mix (RGUs: HSD, CDV, XM, XH, video)
Product Usage (monthly downloads, etc.)
Billing Status (autopay, etc.)
...

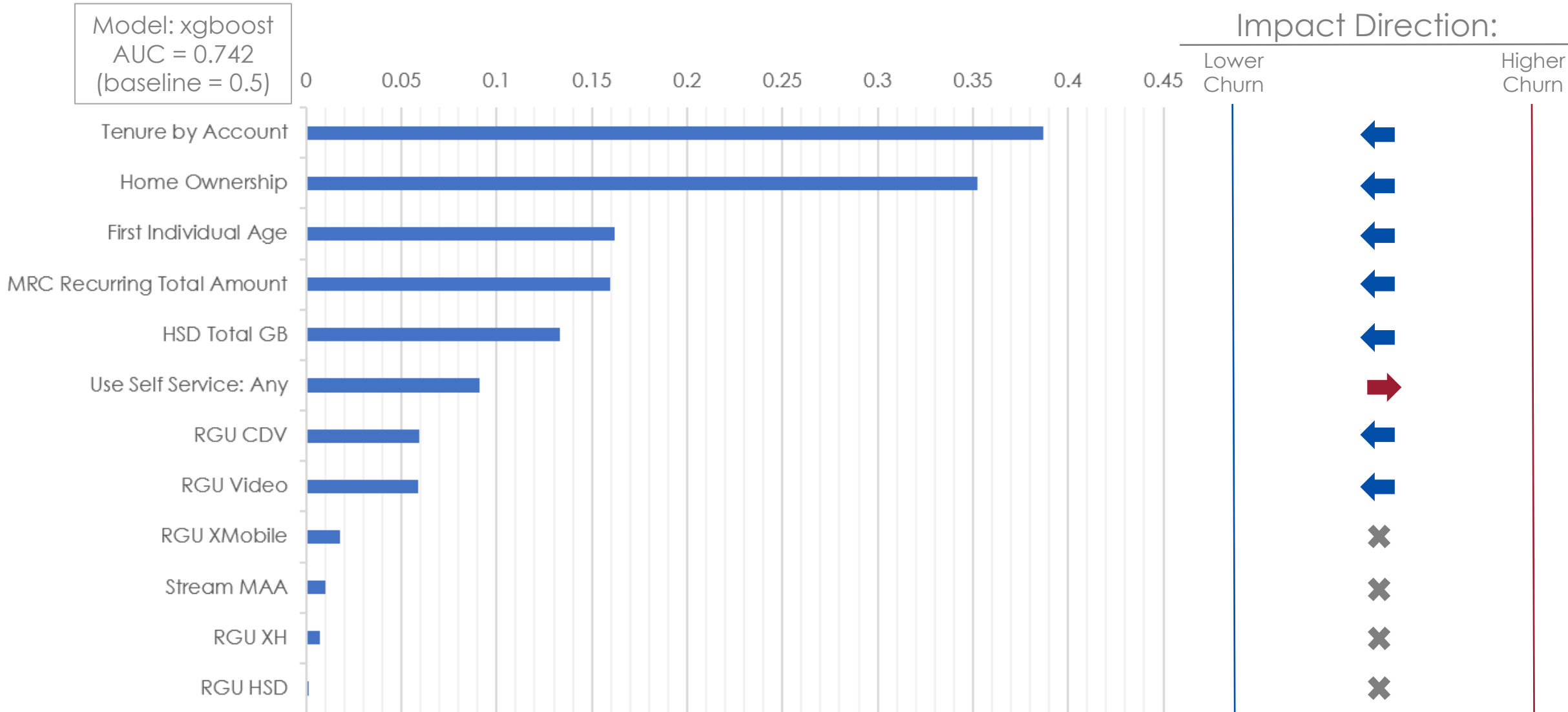
Outcome Variables

Churned during Promo Roll Period
Post-roll status, MRC – trialled but less successfully

APPENDIX 2.2 CHURN RATE DEMOGRAPHIC DRIVERS



Variable Increase
Impact Direction:



APPENDIX 2.3: CHURN PREDICTION BREAKDOWN

Modelling Process:

Initial models with over 100 drivers

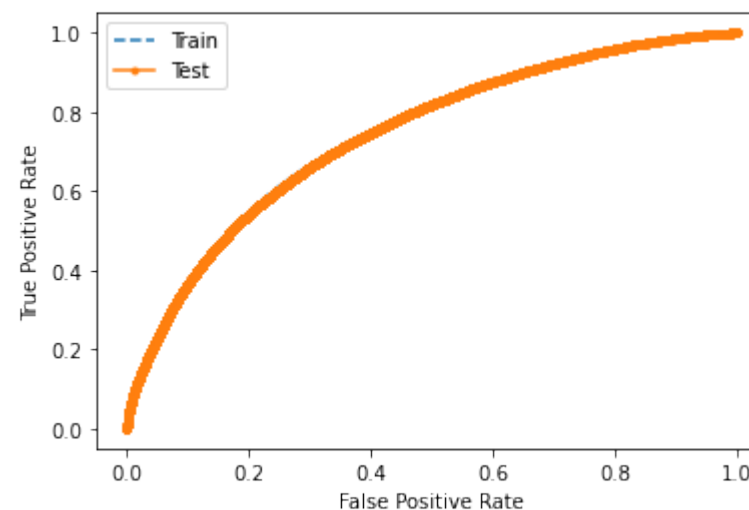
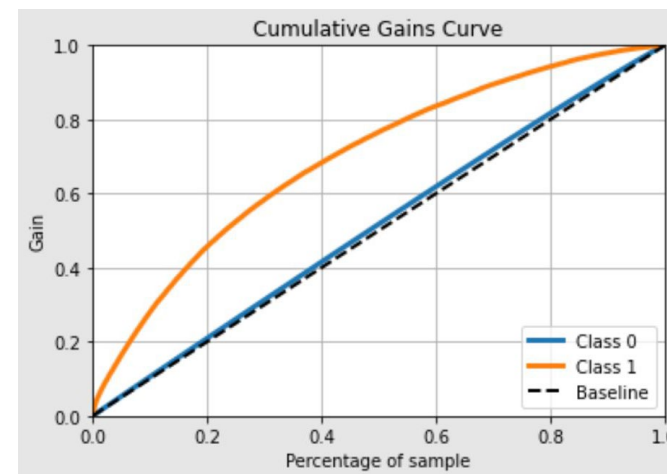
Iterative Model Reduction

Final model with key drivers

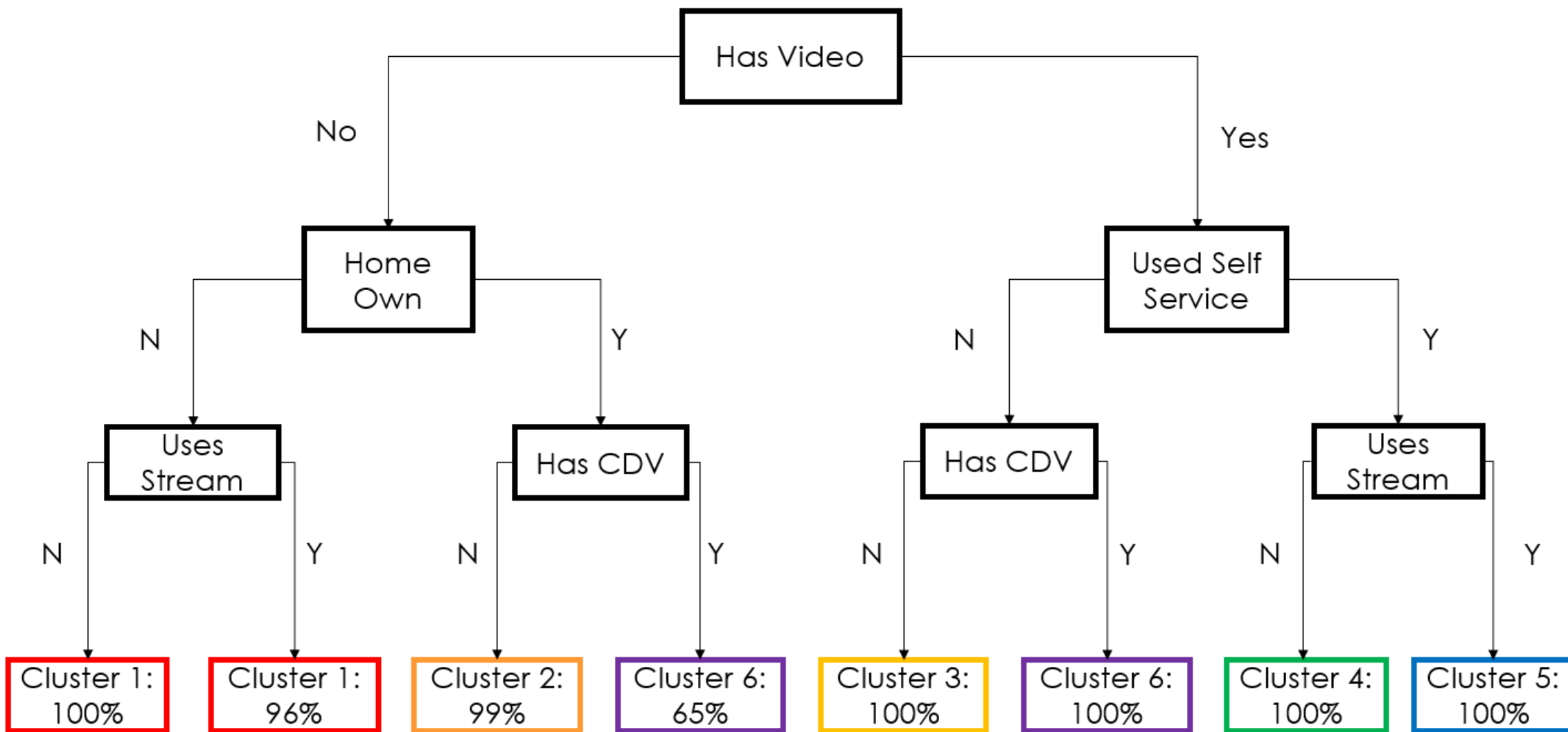
Metrics:

Accuracy – 0.880

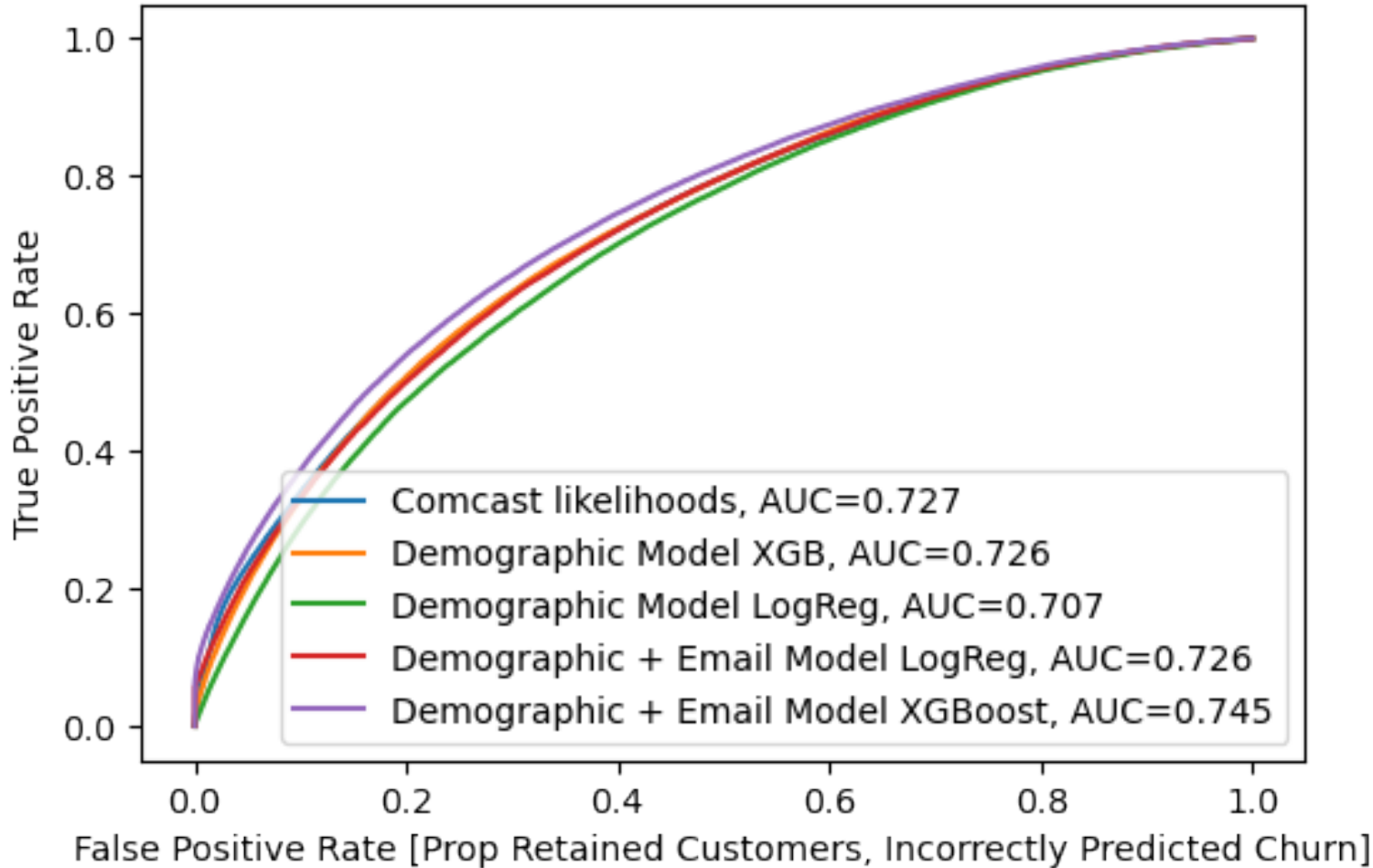
AUC – 0.742



APPENDIX 2.4: FULL CLUSTERING DECISION TREE



APPENDIX 3.1 OVERALL MODEL PERFORMANCE



Curves Displayed:

- Comcast Likelihoods (Comcast)
- Demog. Model (Deliverable 1)
- Demog. Logreg (Deliverable 1)
- Emails_Logreg (Deliverable 3)
- Emails_XGBoost (Deliverable 3)