

# Predicting Subscriber Churn

DAT7, Summer 2015

$$\begin{aligned} &\text{Subscribers} \\ &= \\ &\text{Eyeballs and/or Interactions} \\ &= \\ &\text{Advertising \$\$ \$\$} \\ &= \\ &\text{Profit} \end{aligned}$$

## Problem

Recruiting New Customers Is Difficult:

- Saturated markets mean entirely new subscribers are rare.
- Acquisition is costly
- Product enhancement limitations

## Solution

Retention:

- Existing subscribers are less costly to maintain and serve
- Loyalty – subscribers who stay, interact
- Word of mouth marketing

# Question(s)

- For a given subscriber, what is the likelihood of churn?
- What are the major determinants of churn?

# Data

## Dependent Variable

Subscriberid + Status (U, or S) + Date\_Effective = Churn (binary)

## Independent Variables(?)

Product (200+ in Universe)

Normalized Values (Position Function & Level, Company Type)

Duration of Subscription

Engagement Metrics (Date of Last Open, Date of Last Click)

Type of Acquisition Channel

## Domain Knowledge

Type of Churn

Time

# Modeling

Accuracy is important...

- Why spend extra money on subscribers who would have stayed anyway?

So is comprehensibility...

- Internal stakeholders need to trust that results are consistent with past experiments
- Insight into the determinants of churn (generally)
- Adoption by business users (Audience Development, Sales)

Logistic Regression (Default)

Decision Trees

Support Vector Machines

Survival Analysis (Time to Churn!)