Big Data-Driven Marketing:

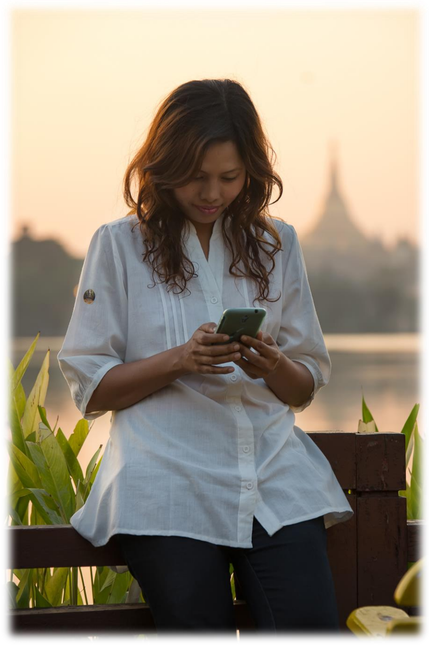
How Machine Learning Outperforms Marketers' Gut-Feeling

Pål Sundsøy[1], Johannes Bjelland[1], Asif M.Iqbal[1], Alex Pentland[2], Yves Alexandre De Montjoye[2]
[1]Telenor Research, Oslo, Norway
[2]Massachusetts Institute of Technology, The Media Laboratory, USA

Abstract

We show that a data-driven approach to text-based marketing outperforms marketers by *13x* in a large-scale experiment in Asia

Using telecom metadata and social network analysis, we created new metrics to identify customers that are the most likely to convert into mobile internet users. These metrics falls into three categories: discretionary income, timing, and social learning.This leads to conversion rates far superior to the current best marketing practices within MNOs.



Best practice is based on experience and gut-feeling

The current best practice in MNOs relies on the marketing team's experience to decide which customers should receive a text for a specific campaign. The marketing team typically selects customers using a few simple metrics directly computed from metadata.
We build a control group based on criteria chosen by the marketing team and includes rules derived from SMS-usage, spending and previous data consumption. Such rules are typical for this type of campaigns.

We predict adoption *and* sustained usage

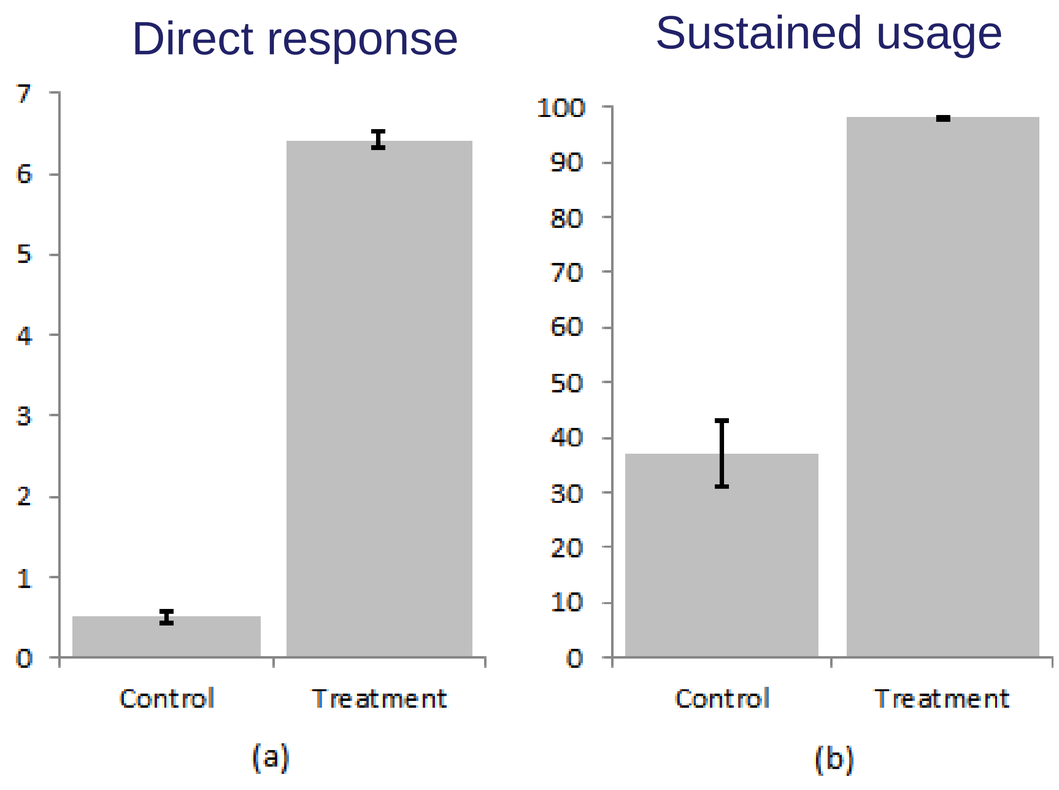
We develop and train the model using 6 months of metadata. Our goal is to identify the behavior of customers who 1) might be interested in using internet and who 2) would keep using mobile internet afterwards. We combine these criteria in the classifier

Social variables are the best predictors

We tested several modeling to classify adoption. The final cross-validated model is a bootstrap aggregated decision tree which performed best on accuracy and stability. The top 10 most useful features to classify natural converters. Ranked by importance in the model.

Rank	Type	Description
1	Social learning	Total spending on data among close social graph neighbors
2	Discretionary income	Average monthly spending on text (binned)
3	Discretionary income	Average monthly number of text sent (binned)
4	Discretionary income	Average monthly spending on value added services over text (binned)
5	Social learning	Average monthly spending on data among social graph neighbors
6		Data enabled handset according to IMEI (Yes/No)
7	Social learning	Data volume among social graph neighbors
8	Social learning	Data volume among close social graph neighbors
9	Timing	Most used handset has changed since last month
10		Amount of 'accidental' data usage

Targeting 250 000 customers gives 13 times better success rate for the data-driven approach



(a) **Conversion rate** in the control (best practice) and treatment (data-driven approach) groups.
(b) The percentage of converted people who **renewed** their data plan after using the volume included in the campaign offer.

Discussion

The success of this pilot study triggered new technical developments and this method is now being put into production.

We expect such an approach will enhance the long-term customer experience by greatly reduce spamming and by providing the customer with more relevant offers.