



HOW TO DECREASE NUMBER OF UNSUBSCRIBERS

Case study in Telecom

October 2015



WWW.DATA-TRACER.COM
MICHAEL.DOPIRA@
DATA-TRACER.COM

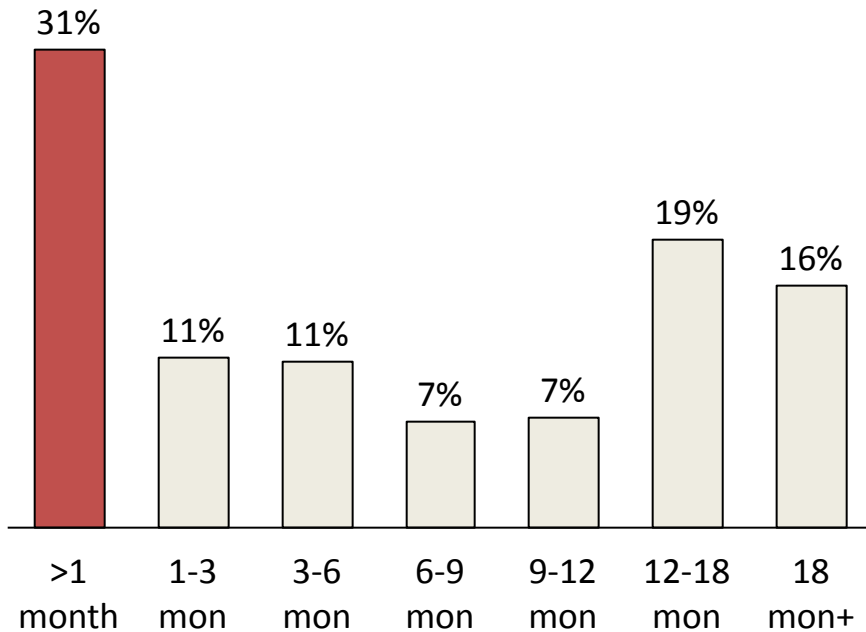
ABOUT THE APPROACH IN BRIEF

- No recommendations based on “expert opinion” or “international best practices”
- All insights are based on deep statistical analysis of the data given by the particular client, so the recommendations are relevant for this client
- Multinomial data analysis techniques are employed. Such approach makes it possible to evaluate the interaction of variables and the importance of each variable
- All models undergo cross-validation, which means that their effectiveness is validated by data that was not used for model building



KEY CHARACTERISTICS OF USERS

“Lifetime” of the clients



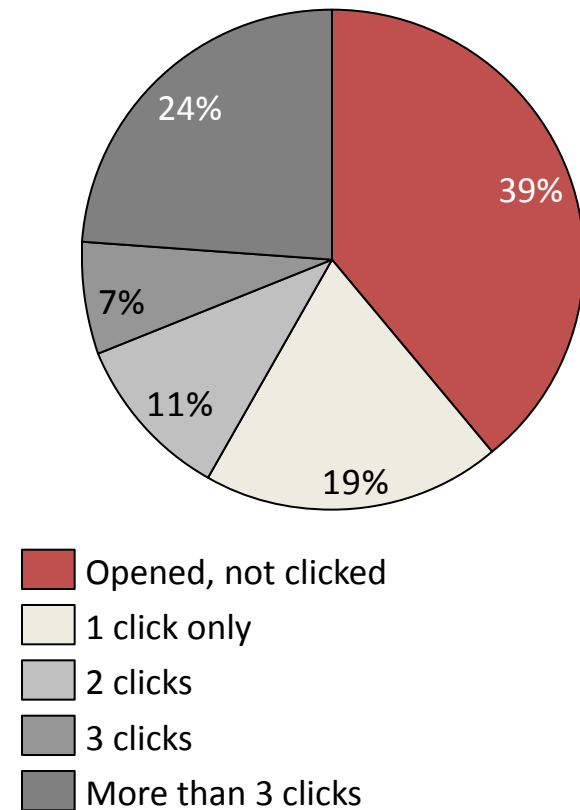
“Lifetime” of 31% of clients is very low – less than 1 month



WWW.DATA-TRACER.COM

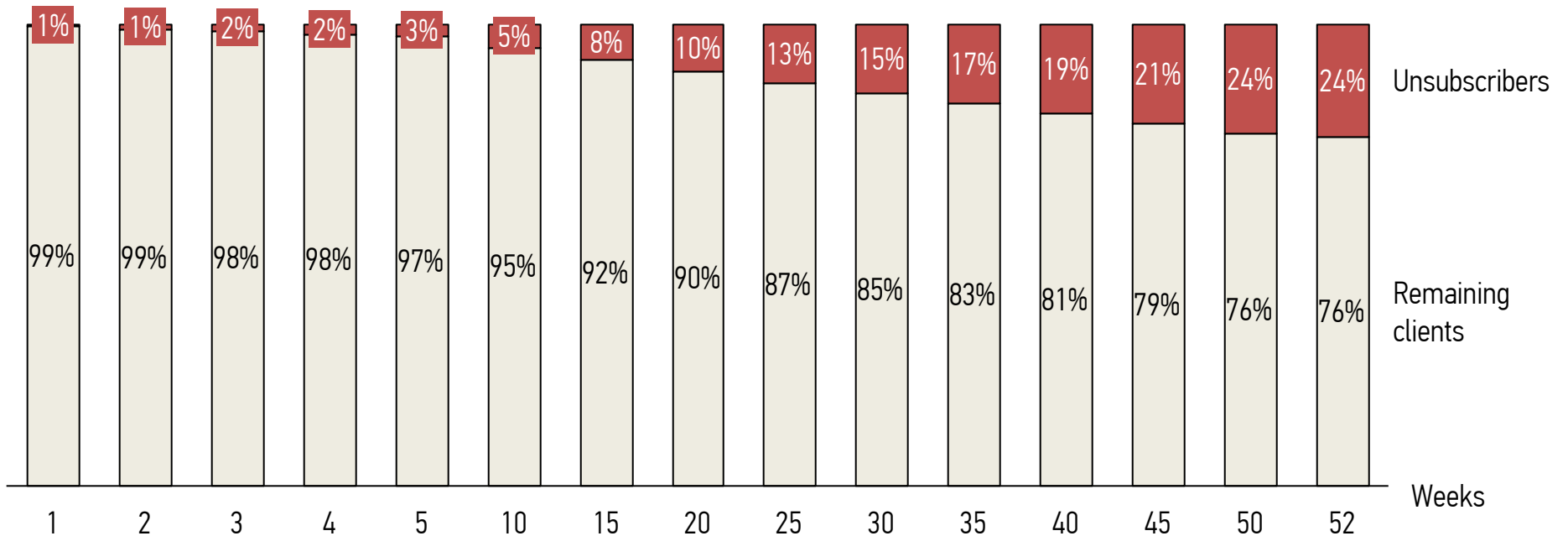
MICHAEL.DOPIRA@DATA-TRACER.COM

Interactions with emails



IF NO NEW USERS ARE ATTRACTED, CUMULATIVE NUMBER OF UNSUBSCRIBERS IS GROWING FAST

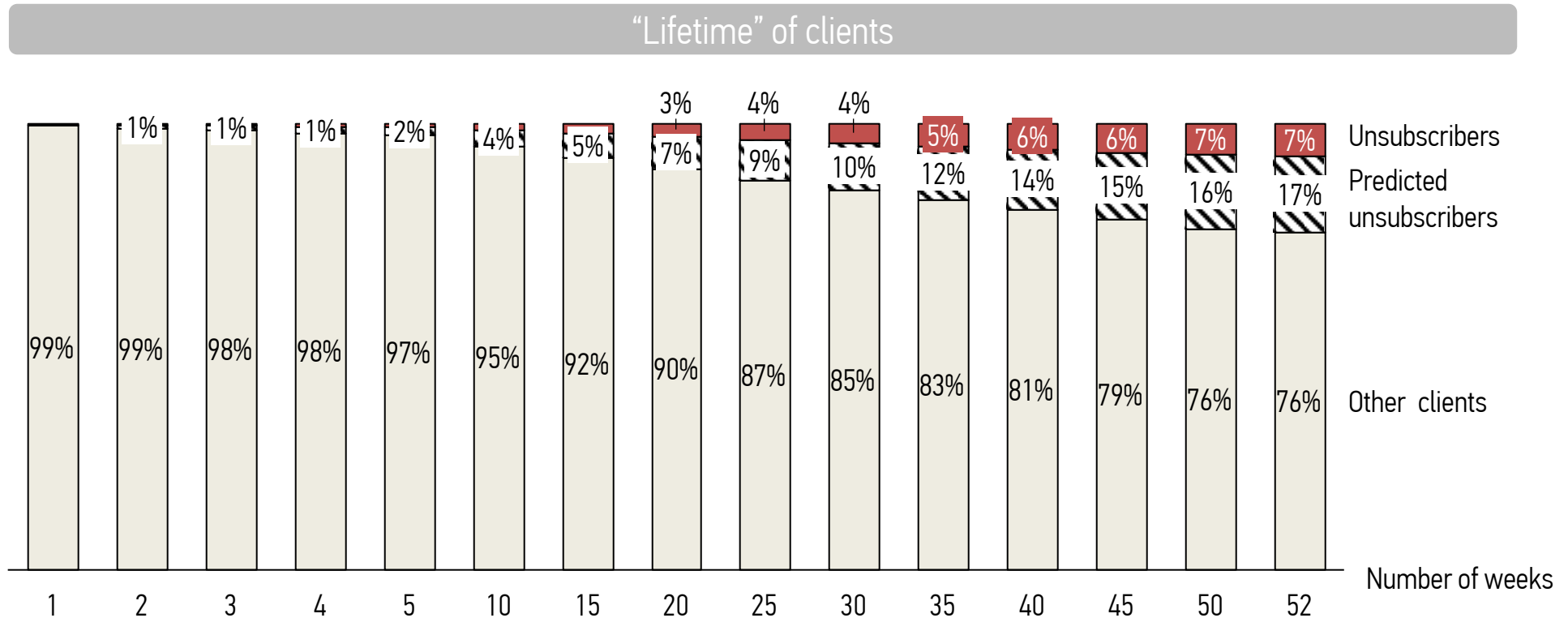
“Lifetime” of clients



About **0.535%** of users unsubscribe every week. If no new users are attracted the yearly churn rate would reach **24%**.



IT IS POSSIBLE TO PREDICT ABOUT 70% OF UNSUBSCRIBERS

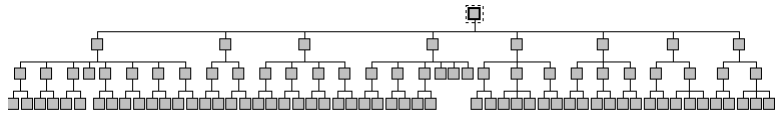


Using results of the analysis it is possible to predict about **70% of potential unsubscribers**, which reduces annual churn rate to **7% only**.



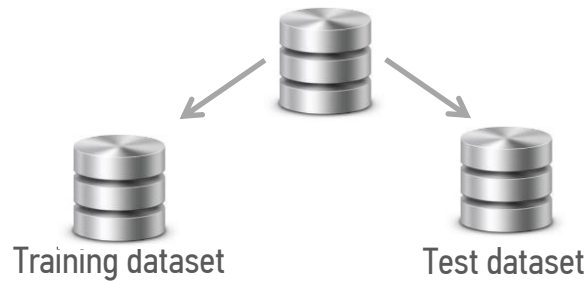
HOW PREDICTIONS ARE MADE

Statistical analysis



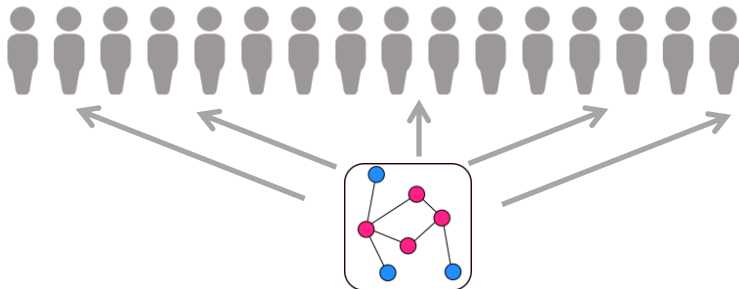
- Using **advanced machine learning tools** (e.g. random forest), it is possible to predict cases when probability of unsubscription increases dramatically.
- For example, if the client has been active for 9 months and starts to use more basic rates, the **probability of unsubscription increases by 63%**.

Validation of the model



- Statistical model is built on one dataset (training dataset), while **model predictions are tested on another dataset (test dataset)**.
- If the model works for the test dataset, it can be **generalized and used for decision making**

Evaluation of probabilities for each user



- Each user is assigned the probability of leaving based on user activity
- Thus, it is easy to pick the users who are under the greatest risk of leaving the company

MODEL CHARACTERISTICS

- Actions of **232,840 active users** were used to build the model
- Training set included **3,281 unsubscribers**.
- Method of analysis – Random Forest.
- **Model correctly predicts over 70% of unsubscribers (AUC - 0.747)**
- Stable predictive power for training and test datasets.





Please contact us if you have any questions:

+38(097)740-6686

+1(307)314-9033

+7(901)903-1791

michael.dopira@data-tracer.com



WWW.DATA-TRACER.COM

MICHAEL.DOPIRA@DATA-TRACER.COM