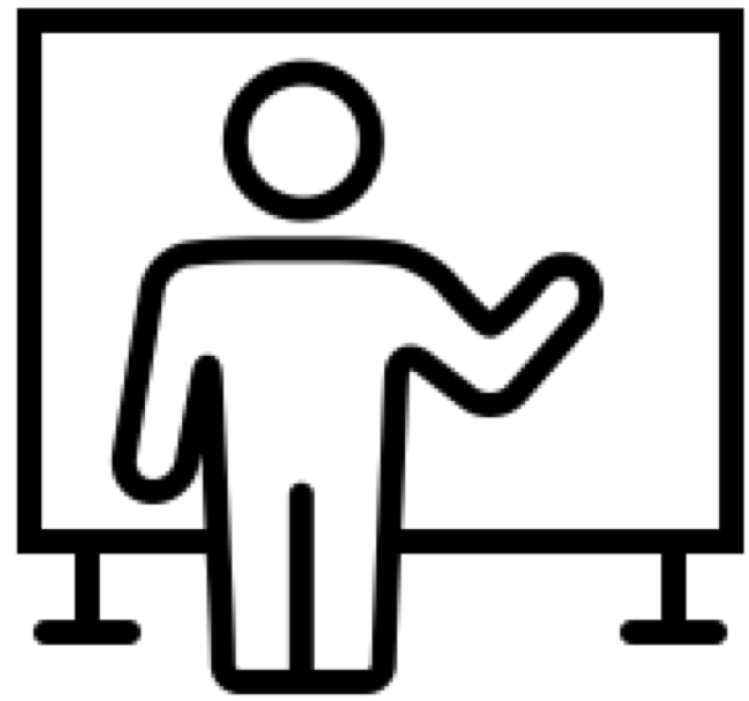


The Comparison of Methods for Individual Treatment Effect Detection



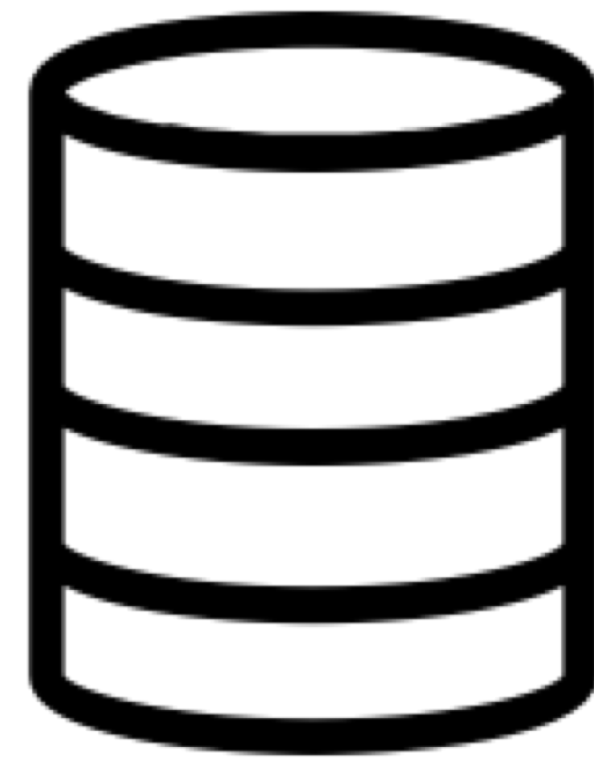
Structure of presentation



Introduction



Approaches of ITE
estimation



Data



Methodology



Results and
conclusion

Introduction



A world of growing information



Data is crucial in many areas of science and business



Particularly, in case of individual treatment effect estimation



Treatment is a certain exposure exerted on an individual or group of individuals in order to provoke a response



Business

effect of push notifications to buy goods at a discount



Medicine

effect of a drug on the human body

Approaches of ITE estimation

$$ITE = E[Y_i = 1 | X_i = 1, treatment = 1] - E[Y_i = 1 | X_i = 1, treatment = 0]$$

1. Two-model approach (Indirect)

Score of ITE is the difference in predictions for models under treatment and not.

(Hansotia & Rukstales, 2002)

2. Modified outcome approach (Direct)

Idea of adding special interaction variables, which become equal to zero on the control group, and on the test group correspond to independent variables of the original data sample.

(Jaskowski & Jaroszewicz, 2012;
Weisberg & Pontes, 2015)

3. Uplift random forest model

(Guelman et al., 2015)

Data



Criteo Uplift Modeling Dataset
(Diemert et. al., 2018)



Belonging to test or control group



25 million observations (users of
the website)



Visited or not the advertiser's website
during the test period (2 weeks)



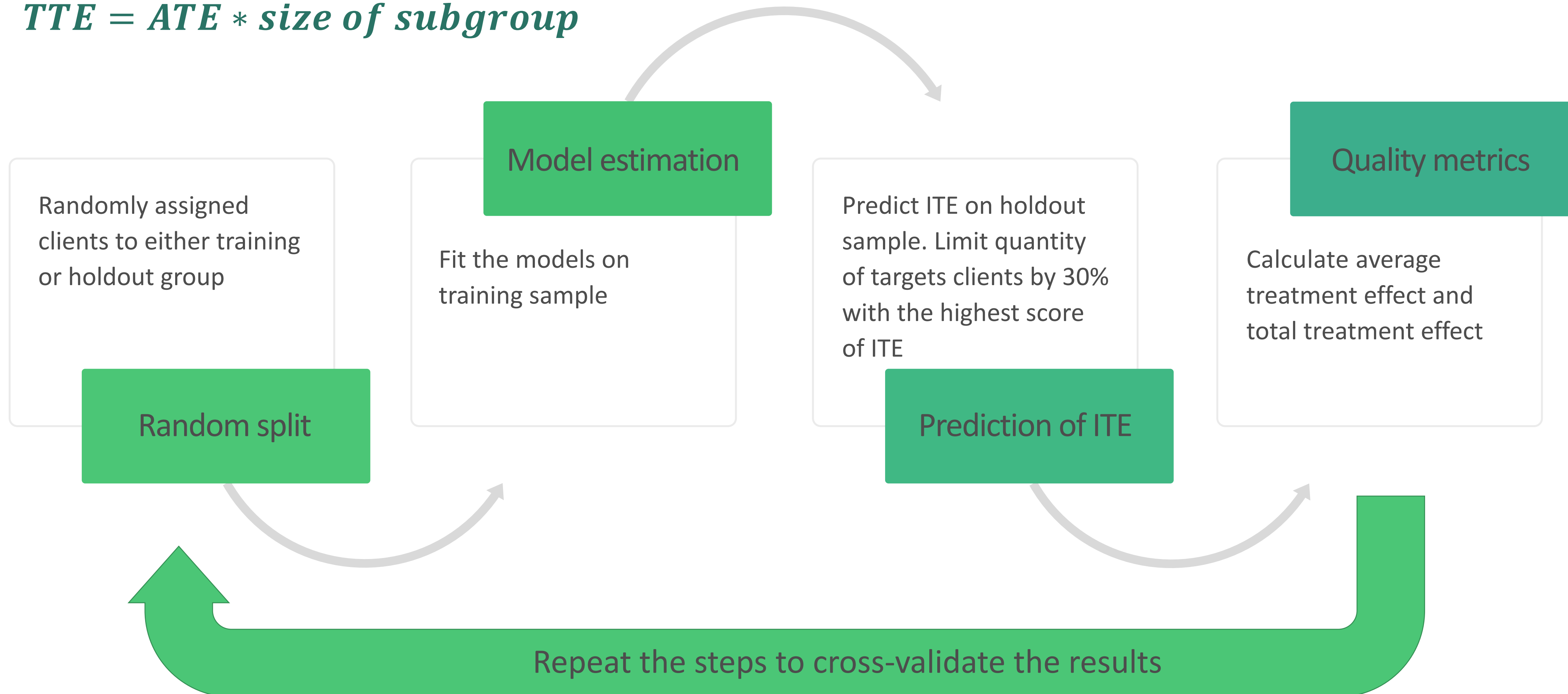
12 features

Methodology

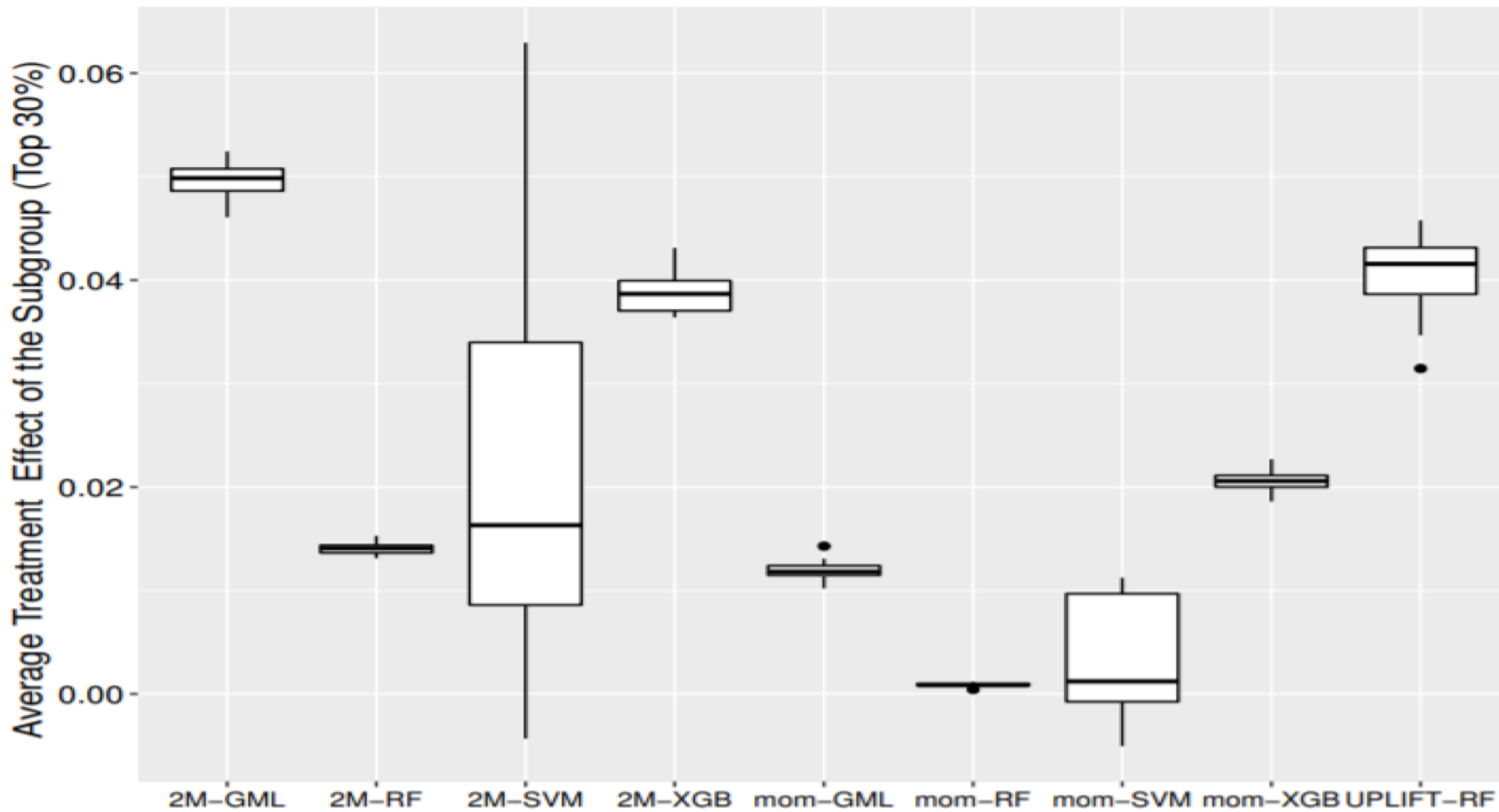
$$ITE = E[Y_i = 1 | X_i = 1, treatment = 1] - E[Y_i = 1 | X_i = 1, treatment = 0]$$

$$ATE = E[Y | treatment = 1] - E[Y | treatment = 0]$$

$$TTE = ATE * size\ of\ subgroup$$



Empirical results



Conclusions



Difference score method estimated by linear logistic regression showed better results.



One of the reasons for the high quality of the Two-model approach could be the particularity of the dataset.



Truncation of train dataset for Uplift Random Forest may be the reason for the poorer performance demonstrated by UPLIFT-RF compared to 2M-GML

The Comparison of Methods for Individual Treatment Effect Detection



Daria Semenova, dvteterina@gmail.com
Maria Temirkaeva, mariya.temirkaeva@mail.ru