

Marketing Strategy

To Treat or Not To Treat? Five Lessons Learned from Using Uplift Modeling to Optimize Marketing Campaigns

Detlef Schoder and Jannik Rößler

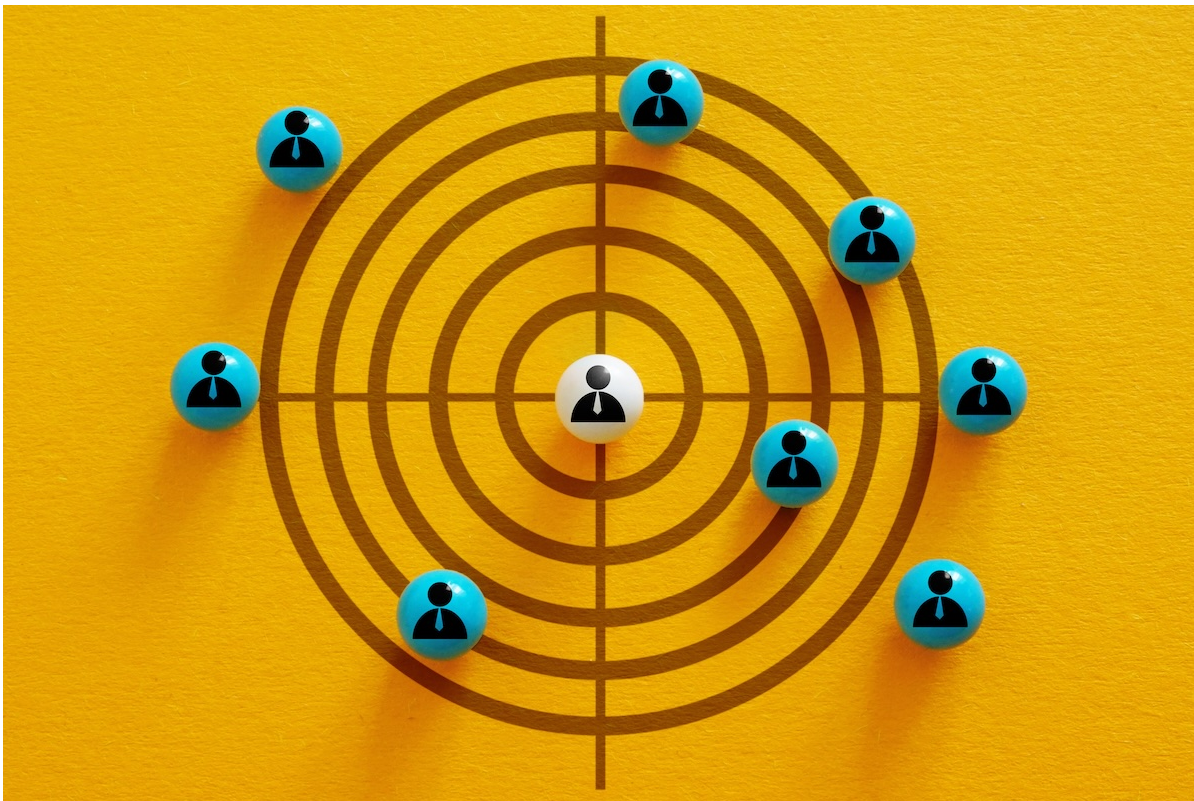


Image Credit | Andrii

Best practices for using an uplift model to increase success and ROI in marketing campaigns.

Conventional wisdom states that online advertisements, retention incentives, sales discounts, and other marketing tools work best when you know and understand your customers. One key task to achieve this understanding is the treatment assignment problem, which focuses on determining the most effective treatment to assign to each individual customer. Assigning the right treatment could persuade a customer to make another purchase; the wrong one could drive a customer away. Gas and electricity suppliers need to send the right promotion to the right households to prevent them from churning. E-commerce platforms need just the right incentives to push customers to buy more products. Retailers succeed when they offer precisely calculated discounts to boost sales.

RELATED ARTICLES

Chaturvedi, Rijul, and Sanjeev Verma. “[Artificial Intelligence-Driven Customer Experience: Overcoming The Challenges](#).” California Management Review Insights, March 1, 2022.

Kumar, V., Bharath Rajan, Rajkumar Venkatesan, and Jim Lecinski. “[Understanding the Role of Artificial Intelligence in Personalized Engagement Marketing](#).” California Management Review 61, no. 4 (2019): 135–55.

RELATED TOPICS

[Predictive Modeling](#)

[Advertising](#)

[Data Analysis](#)

[Business Intelligence](#)

Determining which treatment will work requires modeling the individual treatment effect (ITE). This approach represents the evolution from traditional approaches that either fail to evaluate treatment effects entirely or assess them only in aggregate rather than an individual level (Figure 1).

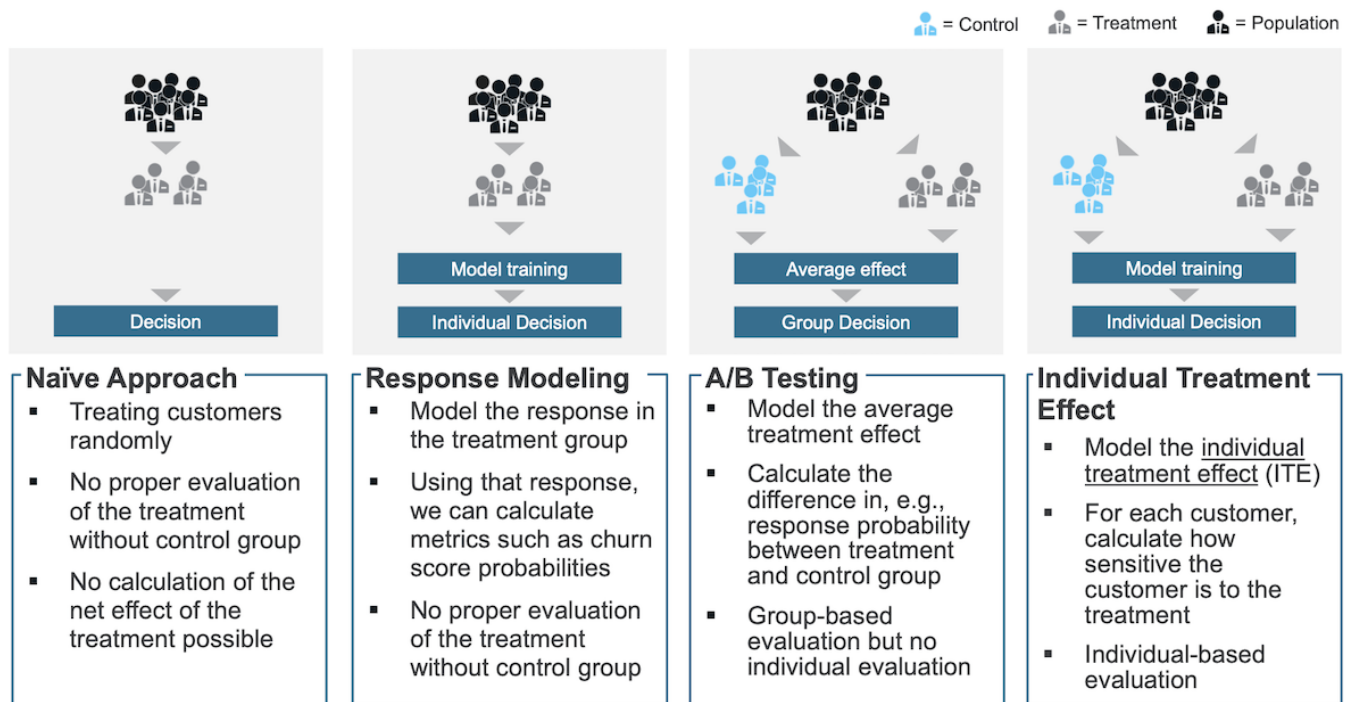


Figure 1: Evolution of treatment assignment methods from traditional approaches to ITE.

But still, postulating the effect of a treatment at an individual level can be really tricky because of the *fundamental problem of causal inference*:¹ we can observe an individual's outcome either after the individual has been subject to a treatment or when the individual has not been subject to a treatment, but never both at the same time. On top of that, there's a cost/benefit problem to consider: the same treatment that increases the likelihood of a purchase may also result in an incremental monetary loss, even in cases where the customer accepts the offer.

Methods that model ITEs are, from a theoretical perspective, far superior because each customer can be assigned to the treatment associated with the most beneficial outcome (e.g., the one with the highest profit).² ³These techniques allow not only *descriptive* and *predictive* analytics—what has happened in the past or could happen in the future—but also *prescriptive* analytics, what should be done today for a desired future outcome. With the

right treatment effect on an individual level, a business can craft campaigns that resonate deeply with target segments, sparking genuine customer engagement. For example, we worked with a retailer to build a model that predicts the ITEs for each customer and increased the conversion rate for the company by 2%. In industries with large customer bases and several campaigns a year, this seemingly small number could translate into millions of dollars in extra profit just through improved design and exploitation of individual (response) data that allows for optimizing marketing campaigns.

Uplift modeling is one of today's hottest approaches to uncovering ITEs—and there are many ways in which it deserves its reputation, from a theoretical perspective, as the best method (Radcliffe 2007; Radcliffe & Surry 1999/2011; Gubela et al. 2019; Baier & Stöcker 2022).^{4 5 6} That said, we've found from our extensive practical experience in a multitude of real-world campaigns that in some situations uplift modeling could yield negative outcomes even if applied in a technically proper way. We also discovered situations in which its complexity and data quality demands make uplift modeling not worth the trouble.

We've developed some key takeaways from our extensive research collaborations as well as many different marketing campaigns in diverse settings: contractual and non-contractual; offline and online; and from promotional campaigns to direct marketing campaigns and advertising. We have what we think is some very strong, practical advice for optimizing large-scale marketing campaigns on an individual level, which we present as *lessons*. We make a particular point of contrasting when ITEs are a good idea and when they fall short.

How Uplift Modeling Works (Briefly)

Uplift modeling predicts the *change* in behavior *caused* by a treatment at an *individual level*. Its implementation begins with a randomized controlled experiment (e.g., by running a pilot campaign) in which customers are randomly assigned to either a *treatment group* – which receives a marketing treatment (e.g., a discount) – or a *control group*, which does not receive any treatment. During the experiment, the company collects the main outcome

variable (i.e., the dependent variable), for example, whether the customer extended his contract or purchased a product, along with demographic, transactional, and psychographic data.

Next, uplift modeling uses both the treatment and control group to build a predictive model that predicts the ITE. Put differently, the model predicts the difference in outcome probability if a customer receives the treatment versus not receiving it. Through these predictive models, uplift modeling makes it possible to distinguish between four customer groups:

- **Sure Things** are customers who exhibit the desired behavior regardless of whether they receive the treatment.
- **Persuadables** exhibit the desired behavior *only if treated*.
- **Sleeping Dogs** exhibit the desired behavior *only if not treated*.
- **Lost Causes** do not exhibit the desired behavior regardless of whether they receive the treatment.

The overall goal of uplift modeling in the ITE context is to identify *Persuadables* and target them for treatment while avoiding doing so for other types of customers (Devriendt et al., 2018).⁷ Only the *Persuadables* segment provides true incremental responses. Traditional response modeling often targets *Sure Things*, lacking the ability to distinguish them from *Persuadables*. As Figure 2 summarizes, *Persuadables* exhibit the desired behavior only if treated, and while that comes with costs, it also generates additional revenue that may be profitable. By contrast, targeting any other type of customers with a treatment mostly just creates additional costs—some more than others.

Once the predictive model has been trained on data from the randomized controlled experiment, it can be used to optimize future campaigns by selectively targeting *Persuadables* and excluding other customer types. The model assigns each customer an uplift score, representing the estimated incremental likelihood that the customer will respond positively to the treatment. The higher the uplift score, the more likely a customer is to be classified as a *Persuadable*. Accordingly, managers and marketing practitioners

can rank customers based on their uplift scores and target the top segment (e.g., the top 10%) in subsequent campaigns, thereby maximizing incremental gains and minimizing the unnecessary costs.

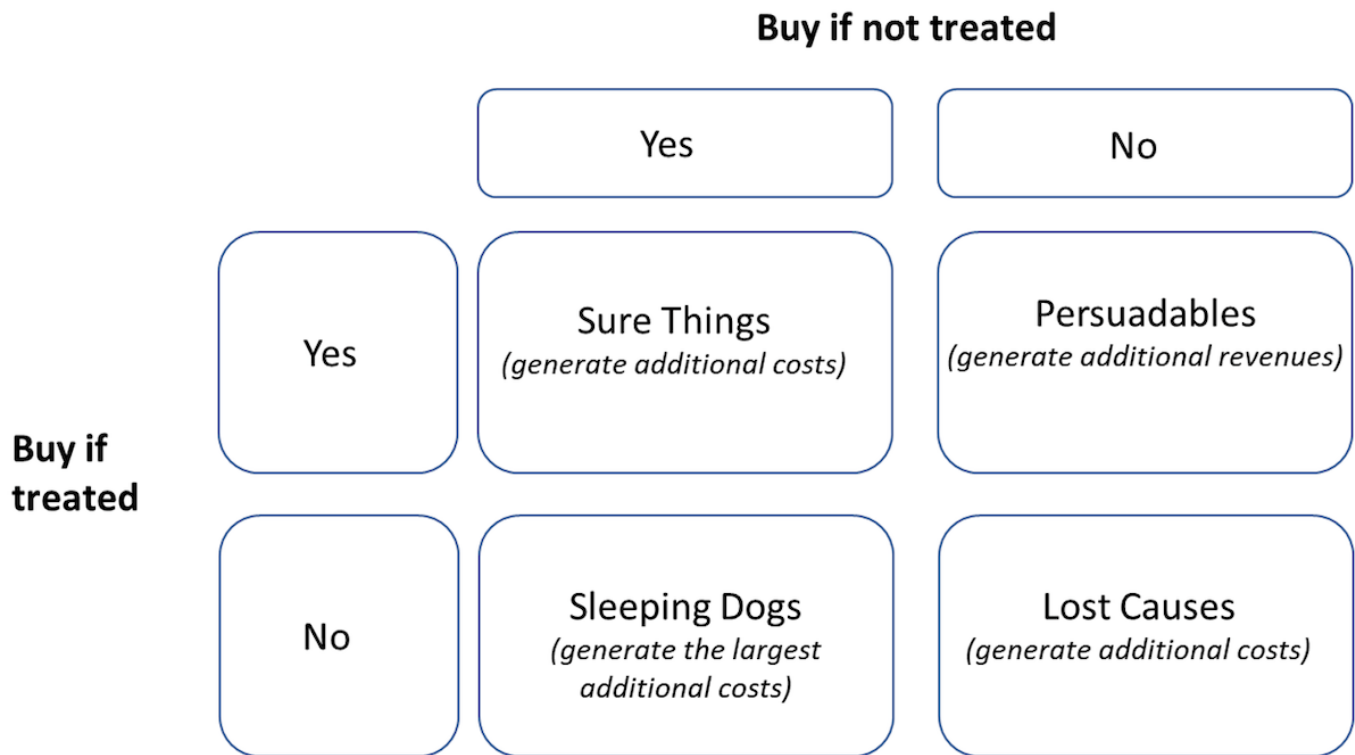


Figure 2: Customer types in marketing campaigns.

Five Lessons for Managers

In a nutshell, uplift modeling offers a way to overcome the problem of not knowing what would have happened if the customer had not been given a treatment. Employing this approach in multiple research and industry collaborations, we've learned five key lessons regarding when using it is worthwhile and when it should probably be avoided—all insights with practical importance for marketing campaigns.

Lesson #1: There are situations when uplift modeling is clearly worth it.

In contractual settings with auto-renewal

Uplift modeling is especially promising in settings where companies have contractual relationships with their customers that include auto-renewal, such as in service agreements (e.g., telecommunication, streaming, software-as-a-service), membership models (e.g., gyms or clubs), and warranty / maintenance contracts (e.g., contractual agreements for maintenance or warranty services in electronics or automotive). The main advantage of uplift modeling in these settings is its ability to identify *Sleeping Dogs*—customers who automatically renew (and thus extend) their contracts if not treated, but who (may) cancel their contracts if they are treated (because the treatment reminds them of the forthcoming auto-renewal). Traditional approaches fail altogether to identify *Sleeping Dogs*, making uplift modeling far superior.

Not looking after *Sleeping Dogs* can be fatal for a marketing campaign. It's futile to target them, the costs incurred are completely unnecessary, and more importantly, targeting these customers actually has a *negative* effect by provoking them to churn.

Depending on the market, there may be very few or even no *Sleeping Dogs*, or this customer type could constitute the largest share, depending on the market. In two of our use cases, one involving a gas/ electricity supplier and the other a telecommunication provider and both operating in low-budget markets, customers frequently switch providers in search of the best prices and thus exhibit the behavior of a *Sleeping Dog*.

When treatment costs are particularly high

Uplift modeling is also superior to traditional methods when the costs for treating customers are particularly high, such as in offline campaigns that target customers through print mailings, telemarketing, and direct door-to-door sales. If you calculate printing costs, postage, the cost of telemarketing staff to make calls, and so on, you can easily see what a colossal waste of money might be involved—an amount that can't possibly be offset by revenue gains even if a treatment changes customer behavior.

The key advantage of uplift modeling in these settings comes from the ability to differentiate between different types of customers and thus drastically reduce the number of them to target, whereas traditional methods simply fail to distinguish *Sure Things* from *Persuadables* and *Lost Causes* from *Sleeping Dogs*.

In one collaboration, we reduced the number of targeted customers by 80%, and the costs of targeting from \$400,000 to \$80,000, while increasing the number of contract renewals compared to targeting all customers (because far fewer *Sleeping Dogs* were bothered).

Lesson #2: You need to figure out how many *Sleeping Dogs* there are.

Determining the proportion of *Sleeping Dogs* in your target market is mainly an empirical question, and can be approximated through (small) test campaigns or by carefully controlling the campaigns you implement.

Research has shown that in many non-contractual marketing settings, *Sleeping Dogs* are either non-existent or there are so few of them that they can be disregarded because they have no (financial) impact.

The absence of *Sleeping Dogs* in a market setting has one important advantage. We can then calculate the upper bound of a marketing campaign, which can be defined simply as the average treatment effect, that is, the difference in conversion rate between the treatment and control groups. Knowing this upper bound helps when evaluating the performance of the uplift modeling algorithms because we can use it to explain how many *Persuadables* a given algorithm correctly identified and to assess whether uplift modeling is even useful (see Lesson #3). In one use case, the conversion rates for the control and treatment groups were 11.04% and 12.21%, respectively, resulting in an average treatment effect of 1.17%—which the relative number of additional purchases due to the treatment can never exceed. Had there been any *Sleeping Dogs*, the upper bound for this retailer could be much higher than the average treatment effect.

We learned from this that with a very high proportion of *Sleeping Dogs*, Uplift modeling is tremendously beneficial and justifies the careful efforts and higher data quality aimed at finding models with really strong predictive power.

Lesson #3: Sometimes it's just not worth optimizing a campaign with uplift modeling.

Under some circumstances, the benefits of uplift modeling may be diminished by various challenges related to data collection and algorithms.

When there are relatively few customers to win and the average treatment effect is small

Let's take the retailer example from above, with the relatively small 1.17% upper bound. If the company has 10,000 customers, optimizing the marketing campaign with any treatment assignment method can at most make 117 more conversions—if the model does a perfect job. Optimizing a model with such a small number of customers—and, accordingly, a small number of likely (additional) customers—may be considered not at all efficient; the overall costs associated with designing an uplift modeling campaign easily trumps the benefits of any treatment assignment, resulting in an unsuccessful marketing campaign. If the company has, say, 100,000+ customers, though, optimizing the marketing campaign with uplift modeling may be a viable option—although decision makers could easily use a less complex method such as response modeling to achieve good results.

When there are few or almost no *Sleeping Dogs*

One of the main benefits of uplift modeling is, as stated, that it can detect *Sleeping Dogs*. Consequently, if there are no or hardly any *Sleeping Dogs*, uplift modeling cannot play out one of its strongest features.

When the treatment is not particularly attractive

Average treatment effects may be small because companies are deploying treatments that are either too small, simply not sufficiently interesting, or just too complex, rendering them not particularly attractive to customers. A 1% discount on the next product purchase would be a case in point. In one of our use cases, the company offered a free switch from one contract to another and ended up inducing a churn rate that was 2% higher in the treatment group than in the control group! We aborted that pilot campaign and introduced another treatment that resulted in a churn rate that was 4% lower, on average, in the treatment group than in the control group.

Interesting, effective treatments could drastically improve the number of conversions, sales, or contract extensions. Companies must, though, make sure their marketing campaigns stay within financial constraints and that the economic value is positive.

When there are almost no costs for treating (prospective) customers

Treatment costs in online settings can be as low as zero. Setting up an online platform can have significant costs, but the incremental cost of reaching additional customers after that can be negligible. In such cases, the benefits of uplift modeling may be overridden—especially if *Lost Causes* constitute the largest share of customers. One company that ran a marketing campaign to acquire new customers had 90% to 95% *Lost Causes*—not unusual in a customer acquisition setup. However, uplift modeling is only superior to traditional methods if treating *Lost Causes* incurs targeting costs. Otherwise, the company can either just treat all prospective customers or use a less-complex method to find the customers with the highest response probability.

When targeting *Sure Things* is perceived as okay

Although targeting them has no effect on *Sure Things* but still incurs costs (including expenses for sending the treatment and the treatment itself), some companies may nevertheless perceive targeting this customer group as okay. They argue that as loyal customers, *Sure Things* need to be kept happy and that their loyalty depends on receiving treatments from time to time.

Lesson #4: Scaling is really important.

Frequency of marketing campaigns

In general, the more marketing campaigns you're running throughout the year, the more important it is to use sophisticated treatment assignment methods. Let's return to the retailer with the 1.17% upper bound for more conversions: assume the company has 50,000 customers and that each additional conversion results in a net revenue of \$50. In the best-case scenario, the company could obtain 585 additional conversions, which yields

an additional \$29,250 net revenue for a single campaign. Run such a marketing campaign twice a year, and the additional revenue is \$58,500; six times a year, it comes to \$175,500; and if done monthly basis, the additional revenue is \$351,000.

Lesson #5: You must choose between focusing on conversion rates or economic value.

The conventional wisdom holds that the more conversions, or more responders, or more contract extensions, the more successful the marketing campaign. That's faulty thinking, though, unless that success serves an ultimate, *explicitly economic* goal: maximize profits given a particular marketing budget, or minimize the needed marketing budget given a particular profit goal.

In one of our collaborations, for example, conversion rates were of 15.07% in control group, 18.04% in the treatment group, and 17.32% in the uplift modeling optimized group. Those numbers suggest that targeting all customers (i.e., the treatment group) would be the best way to optimize the given campaign. Looking at profit per customer for each group (i.e., considering treatment costs and profit margin), however, the *economic* achievement is rather different: \$5.26 profit per customers in the control group, \$5.18 in the treatment group, and \$5.46 in the uplift modeling optimized group.

What This Means for Managers

Managers, whether in contractual or non-contractual settings, can begin optimizing their campaigns by answering the following questions:

- What are the shares of *Sleeping Dogs*, *Persuadables*, *Lost Causes*, and *Sure Things* in our market?
- Are we operating in a low-budget market (i.e., do we need to be aware of *Sleeping Dogs*)?
- Can we design genuinely attractive treatments at lower cost?
- What is the ultimate purpose of our marketing campaign (e.g., CRM, acquisition, conversion)?

While uplift modeling is not always required for designing campaigns, a look at best practices shows that its use can increase the likelihood of success, improve return on investment, and help leaders get the most out of their campaigns.

References

1. Donald B. Rubin. “**Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.**” *Journal of Educational Psychology* 66/5 (1974): 688–701.
2. Kathleen Kane, Victor S.Y. Lo & Jane Zheng. “**Mining for the Truly Responsive Customers and Prospects Using True-Lift Modeling: Comparison of New and Existing Methods.**” *Journal of Marketing Analytics*, 2/4 (2014): 218–238.
3. Jannik Rössler and Detlef Schoder. “**Bridging the Gap: A Systematic Benchmarking of Uplift Modeling and Heterogeneous Treatment Effects Methods.**” *Journal of Interactive Marketing*, August 11, 2022. .
4. Nicholas J. Radcliffe. “Using Control Groups to Target on Predicted Lift: Building and Assessing Uplift Models,” *Direct Marketing Analytics Journal* 1 (2007): 14–21.

Nicholas J. Radcliffe and Patrick D. Surry. “Differential Response Analysis: Modeling True Response by Isolating the Effect of a Single Action,” in *Proceedings of Credit Scoring and Credit Control VI*. (Credit Research Center, University of Edinburgh Management School, 1999).

Nicholas J. Radcliffe and Patrick D. Surry. “Real-World Uplift Modelling with Significance-Based Uplift Trees,” in *White Paper TR-2011-1*, (Stochastic Solutions, 1–33, 2011).

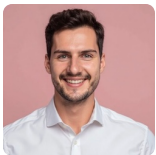
5. Robin Gubela, Artem Bequé, Stefan Lessmann, and Fabian Gebert. “**Conversion Uplift in E-Commerce: A Systematic Benchmark of Modeling Strategies.**” *International Journal of Information Technology & Decision Making*, 18/3 (2019): 747–791.
6. Daniel Baier and Benedikt Stöcker. “**Profit Uplift Modeling for Direct Marketing Campaigns: Approaches and Applications for Online Shops.**” *Journal of Business Economics*, 92 (2022): 645–673.

7. Floris Devriendt, Darie Moldovan, and Wouter Verbeke. “A **Literature Survey and Experimental Evaluation of the State-of-the-Art in Uplift Modeling: A Stepping Stone Toward the Development of Prescriptive Analytics.**” *Big Data*, 6/1 (2018): 13–41.
-



Detlef Schoder [Follow](#)

Detlef Schoder is Chaired Professor of Information Systems & Information Management, serving also as Founding Director of the Cologne Institute for Information Systems (CIIS) at the University of Cologne, Germany. His research focuses on digital transformation and innovation, AI for public communication, protein-binding research, crypto/blockchain, and financial trading.



Jannik Rößler [Follow](#)

Jannik Rößler received his B.S. and M.S. degrees in Information Systems from the University of Cologne in 2017 and 2019 and his Ph.D. in 2023. He is Co-Founder of Pixit, leading generative AI model development in computer vision. His research interests include machine learning, causal inference, and computer vision.