Uplift modeling using the Transformed Outcome Approach

Paulo Pinheiro[1[0000-0002-8912-2244]] and Luís Cavique[2[0000-0002-5590-1493]]

[1] Universidade Aberta and Cedis, Lisboa, Portugal
[2] Universidade Aberta, LASIGE, Lisboa, Portugal
ppinheiro@cedis.pt, luis.cavique@uab.pt

Abstract. Churn and how to deal with it is an essential issue in the telecommunications sector. Within the scope of actionable knowledge, we argue that it is crucial to find effective personalized interventions that can lead to a reduction in dropouts and that, at the same time, make it possible to determine the causal effect of these interventions. Considering an intervention that encourages clients to opt for a longer-term contract for benefits, we used Uplift modeling and the Transformed Outcome Approach as a machine learning-based technique for individual-level prediction. The result is actionable profiles of persuadable customers that increase retention and strike the right balance between the campaign budget.

Keywords. Uplift Modelling, Causal Effect, CATE, Decision Trees, Transformed Outcome Approach

1 Introduction

It is widely accepted that it is a good idea to send a marketing offer to all of its customers, thinking that the customer's likelihood of purchasing a product or service increases. Likewise, when we approach the churn problem, we think that customers identified as most at risk of churn should all be targeted by proactive retention programs.

Some authors [1][2] argue that since customers respond differently to retention interventions, companies should not seek to intervene with all those at risk of giving up or canceling their subscription but only with those who are more sensitive to the intervention. In this sense, we want to know if a customer will maintain the service if we intervene with him or if he will maintain the service even if we do not intervene. Specifically, we want to know whether or not an intervention influences a client.

In marketing, as in statistics, biomedicine, and other areas, knowing whether and to what extent the value of one variable (the treatment) affects the value of another variable (the outcome) is a crucial issue. This question is a problem in causal inference where two paradigms are known: the paradigm of causal structural models [3] and the potential outcome framework [4].

The paradigm of causal structural models begins by noting that a variable X is a direct cause of the variable Y if X appears in the function that affects a value to Y (Y =f(X)). However, the intervention/treatment T must include a new type of variable in the causality task. In this task, the outcome Y of treatment T is the object of study. For this purpose, test and control data sets are used for treatment performed T=1 and not performed T=0. In analogy with f(X, Y), the explanatory function uses three variables, f(T, X, Y). This dichotomy can be found in Pearl and Mackenzie's ladder of causation [5] and the work of Hernán et al. [6].

In this sense, to answer the causal inference question, we need to estimate the causal effect of treatment on outcome. In this work, we use methods of Uplift Modeling that are implicitly designed for data from randomized experiments.

There are several methods to apply Uplift Modelling and find the causal effect we will refer to in related work. This work shows that it is possible to use Uplift Modeling to obtain actionable knowledge concerning finding the consumer profiles most likely to keep their service if subjected to intervention and measure the causal effect of that intervention and its impact on the business. Specifically, based on Telco's dataset, we consider a possible intervention to change the duration of the contracts to find the profile of the customers most likely to maintain the service to maximize return on marketing investment.

The paper is organized into five additional sections. In Section 2, related work is presented. Section 3 details the Telco case study. Section 4 discusses the profiles found and those that can be persuadable. In Section 5, conclusions are drawn and points out some points that require further investigation and work.

2 Related Work

2.1 What is Uplift?

It is common in marketing to find predictive models to understand which customers will buy if they are intervened. Uplift Modeling seeks to know if the customer will buy only if he is intervened. The difference lies in the fact that in Uplift Modeling, we can predict which customers can be persuaded and who reacts because of an intervention. Siegel [1] points to the Uplift Modeling to predict the influence on an individual's behavior that results from applying one form of treatment over another.

In order to learn to distinguish influenceable clients, those to whom it makes a difference to perform some treatment, the Uplift Model learns from both types of clients, those who were contacted and those who were not, so it is necessary to use two sets of data for training the model. The first group of clients that are intervened or treated, the treatment group, and the second group of clients that do not receive the treatment, is named the control group. Each customer is classified into one of four quadrants, as shown in Fig. 1.

The resulting matrix is created based on the customer's decision to buy (or keep the service in case of a churn problem) depending on whether or not a marketing campaign targeted it. The four quadrants refer to:

- *persuadable* customers who buy when exposed to the marketing campaign (in healthcare are known as *compliers*);
- to secure customers (*sure things*) who buy regardless of whether or not they are the target of a treatment;
- the *lost causes* are customers who, regardless of whether or not they are the target of treatment, will not buy a product or service;
- moreover, the group of clients who should not be treated (*do-not-disturbs*, also known as *sleeping dogs*) under penalty of becoming dropouts.

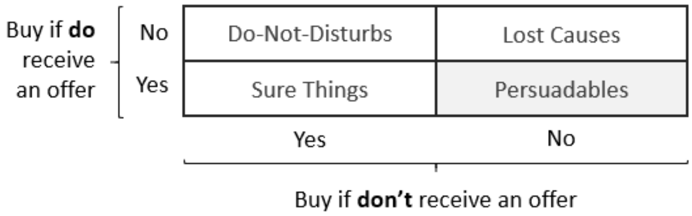

Fig. 1. Conceptual response segments (adapted from [1]).

The aim of Uplift Modeling is thus to identify persuadable customers and avoid treating customers classified as do-not-disturb.

Devrient *et al.* [7] add that Uplift modeling focuses on decision making at the individual level and tries to estimate the causal effect of a treatment on an outcome, thus determining which treatment to apply to each individual to optimize the outcome. On the other hand, it allows determining the success of a campaign by observing customers' behavior in the treatment and control groups in terms of response. It allows the calculation and comparison of the response rate for both groups, and the difference in response rate is the increase due to the campaign.

2.2 Uplift problem

As mentioned before and illustrated in Fig. 1, the goal of Uplift Modeling is to find persuasive customers and not intervene with do-not-disturb ones. However, the data we have available for building models only contain information on whether or not the client intervened and whether or not he responded, as illustrated on the left side of Fig. 2. The problem that arises is how to frame customers as intended by the Uplift Modeling.

As summarized by Devrient *et al.* [7], given the data obtained in previous interventions, we can group customers into four categories: those who responded and were intervened (TR); in this case, we do not know if they would have responded even if they had not been intervened, reason why they can be *persuadable* or *sure-things*; those who did not respond and were intervened (TN); in this case, we do not know if they would not have responded if they had not been intervened so that they could be *lost causes* or *do-not-disturbs*; those who responded without being intervened (CR); in this case, we do not know if they would also have responded if they had been intervened, so we do not know if they are *sure-things* or *do-not-disturbs*; those who did not respond and were not intervened (CN); in this case, we do not know if they would have responded if they were intervened, the reason why they can be *persuadable* or *lost-causes*.

As it is not expressly possible to separate only persuadable customers, it is not a significant risk to intervene between the categories of those who responded when intervened (TR) and those who were not intervened and did not respond (CN). Although these categories include persuadable clients, lost causes, and sure-things that should not be dealt with, their intervention involves a lower cost than dealing with do-not-disturbs.

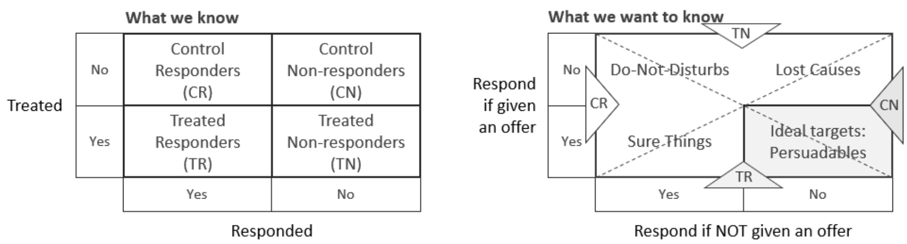

Fig. 2. Conceptual table of what we know and what we want to know (adapted from [7])

After applying techniques that seek to fit clients into these categories, it is possible to find the measure of the Uplift. Ascarza [2] has addressed this problem, then states that the clients to be intervened, the persuadable ones, will be those who have an Uplift measurement above a minimum effect to consider.

2.3 Measuring Uplift

As mentioned by Gutierrez & Gérardy [8], assuming $Y_i(1)$ as the result of person i when he is intervened and $Y_i(0)$ as the result of person i when he is not, the causal effect can be given by:

$$\tau_i = Y_i(1) - Y_i(0) \tag{1}$$

The expected causal effect of an intervention in a subgroup of the population, called Conditional Average Treatment Effect (CATE), is given by:

$$\tau(X_i) = E[Y_i(1)|X_i] - E[Y_i(0)|X_i] \tag{2}$$

where X_i is a vector of random variables (features) that describe the pre-intervention characteristics of the individual. $Y_i(1)$ and $Y_i(0)$ are never observed at the same time because a person *i* cannot at the same time be and not be intervened. Thus, considering a binary variable T_i that takes the value 1 if the person *i* is intervened and 0 otherwise, we can write that the observed result Y_i^{obs} is in fact:

$$Y_i^{obs} = T_i Y_i(1) + (1 - T_i) Y_i(0) \tag{3}$$

Assuming that T_i conditional on X_i is independent of $Y_i(1)$ and of $Y_i(0)$ an assumption stated by Rosenbaum & Rubin [9] known as Unconfoundedness

Assumption or Conditional Independence Assumption (CIA), expressed by:

$$\{Y_i(1), Y_i(0)\} \perp\!\!\!\perp T_i | X_i \tag{4}$$

we can then consider estimating CATE from observational data by computing:

$$\tau_i = E\left[Y_i^{obs} \middle| X_i, T_i(1)\right] - E\left[Y_i^{obs} \middle| X_i, T_i(0)\right] \tag{5}$$

Zhang et al. [10] mentioned that since Uplift modeling techniques assume that data are obtained from experiments with randomized treatment assignment, the Uplift calculation is equivalent to the CATE calculation as expressed in equation (5).

2.4 Uplift techniques and the Transformed Outcome approach

Siegel [1] refers to the persuasion effect as the effect obtained in an individual's persuasion by combining the paradigms of comparison of the results obtained in treatment and control groups and applying predictive modeling, namely through machine learning, statistical regression, and other techniques.

There are several approaches to Uplift Modeling through Machine Learning. There are proposals for classifying the methods according to how they approach Uplift Modeling in the literature review. Gutierrez & Gérardy [8] mentions the two-model approach, class transformation, and Uplift direct modeling. Devriendt *et al.* [7] classify them into two large groups: pre-processing data approaches that include transformation approaches, variable selection procedures, net weights of evidence, and net information value; and the data processing approach in which the two-model approach and direct estimation approaches are framed; Finally, Zhang *et al.* [10] mentions methods using existing supervised learning models, and specific methods for Uplift modeling.

In approaches that use existing supervised learning models, we find the S-Learner, T-Learner, X-Leaner, and R-Learner methods based on Deep Learning and the Transformed Outcome. The specific methods for Uplift modeling include approaches based on Decision Trees, SVM, Deep Learning, and Ensemble Methods.

In this work, we will use the Transformed Outcome Approach. We can find this approach in works such as Athey & Imbens [11] and Jaskowski & Jaroszewicz [12]. It lies in transforming the observed result Y into Y^* such that the CATE is equal to the conditional expectation of the transformed result Y^*.

According to Athey & Imbens [11], verifying the Unconfoundedness assumption referred to before in expression (4), the transformed result Y_i^* can be expressed by:

$$Y_i^* = T\frac{Y_i}{\mathrm{e}(x)} - (1-T)\frac{Y_i}{(1-\mathrm{e}(x))} \tag{6}$$

where $\mathrm{e}(x)$ is the propensity score defined as:

$$\mathrm{e}(x) = P(T = 1|x) \tag{7}$$

moreover, the CATE can be expressed by:

$$\tau(x) = E[Y_i^*|X = x] \tag{8}$$

thus, any supervised algorithm that uses Y^* as the target and X as features can be used.

For binary results, Jaskowski & Jaroszewicz [12] propose a transformation of the outcome where $Y^* = 1$ in the cases where (T = 1 and Y = 1) and where (T = 0 and Y = 0), and $Y^* = 0$ in all other cases, which results from the following equation (9).

$$Y_i^* = Y_i T + (1 - Y_i)(1 - T) \tag{9}$$

It corresponds to the intervention groups mentioned in Section 2.2, which we can see on the right in Fig. 2. Considering that Y is binary and that $e(x) = 0.5$ because an individual has an equal probability of being in the treatment or control group, these authors proved that, in this case, CATE could be estimated by transforming equation (6) which results in:

$$\tau_{(x)} = 2P(Y_i^* = 1|x) - 1 \tag{10}$$

As pointed out in several literature reviews on Uplift modeling [7][8][10], the Transform Outcome Approach is simple. It tends to obtain better results than other approaches based on supervised methods, namely the two-model one. In addition, it has the flexibility to use any supervised method and to calculate the CATE directly. However, they point out as restrictions the dependence of a precise estimate of the propensity score in the case of continuous results. In the case of binary results, the balance between the treatment and control data set.

2.5 Evaluation metrics

It is a common opinion found in literature reviews [7] [8] [10] that the evaluation metrics of traditional predictive models are not adequate to evaluate Uplift modeling because we can never observe at the same time the effect of intervening or not intervening in an individual. The authors present several metrics that can be used, such as the Gini and Qini coefficients. Devriendt *et al.* [7] also refer to the determination of precision in the estimation of heterogeneous effects (PEHE) and mean absolute percentage error (MAPE) as metrics that can be used with known ground-truth treatment effects.

Gutierrez & Gérardy [8] also prove that, in the case of the transformed outcome approach, the Mean Square Error (MSE) (previously referred to by PEHE), in the form:

$$MSE(\tau_i, \hat{\tau}_i) = \frac{1}{n} \sum_i^n (\tau_i - \hat{\tau}_i)^2 \tag{11}$$

can be approximated by:

$$MSE(Y_i^*, \hat{\tau}_i) = \frac{1}{n} \sum_i^n (Y_i^* - \hat{\tau}_i)^2 \tag{12}$$

Another more visual way of analyzing the quality of the Uplift model is through the use of the Uplift chart, described by Devriendt *et al.* [7]. After building the model, the individuals in the training dataset are scored according to the corresponding uplift value, sorted in ascending order of this value, and grouped into deciles. The response rate increment in each decile is calculated by subtracting the response rate of subjects in the control group from the response rate of subjects in the intervention group. In an ideal model, the graph obtained is similar to that in Fig. 3, where the persuasive individuals will appear in the leftmost deciles and have higher uplift values. However, as they point out, these graphs are never presented so regularly in practice.

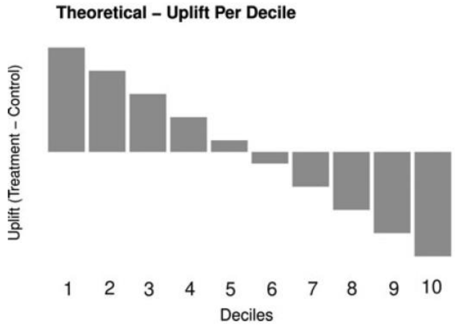

Fig. 3. Uplift chart from an ideal Uplift Model [7]

3 Finding Uplift

This case study presents the Telco Customer Churn public dataset [13] that contains information on eighteen covariates potentially related to the outcomes of interest (churn or not churn). Our goal is to determine which customers can be influenced by the proposal to change the duration of the contract and the proposed intervention to avoid abandoning or canceling the service. With this in mind, the operations are carried out on Procedure 1.

Procedure 1: Uplift modelling

1. Given a dataset D, define the result Y, the intervention/treatment T to be considered, and the set of available covariates X;
2. After balancing the dataset D, the transformed outcome Y* is obtained, and the feature importance is observed through net information value;
3. After splitting dataset D into train D_t and prediction D_p groups, decision tree algorithm is applied to the training dataset D_t, and a decision tree is obtained. The CATE/Uplift is calculated for each of the nodes and leaves of the tree. The rules that avoid churn can be extracted from the leaf node that have higher CATE/Uplift;
4. In other to calculate performance measures, the model is evaluated by using the prediction dataset D_p.

Finally, we extract the rules with the best Uplift score to identify the persuadable customers.

3.1 Outcome (Y) and intervention (T) definition

Each row represents a customer in this dataset, and each column contains the customer's attributes. Those attributes can be grouped in demographic info about customers, attributes that describe customer's account information and attributes that present the services that each customer has signed up for. There is also an attribute, Churn, which indicates whether or not the customer has abandoned services in the last month.

Table 1 shows the number of customers who abandon and do not abandon the service depending on the duration of their contract. As we can see, the percentage of customers who abandon the service when they have a month-to-month contract (42,71%) is substantially higher than the percentage of customers who abandon it when they have contracts of longer duration (7,24%).

Considering that the Contract attribute is actionable, it can target an intervention to make customers switch to long-term contracts to reduce churn.

Table 1. Number of customers that churn by Contract duration

Contract attribute class	Churn attribute		Total
	Yes	No	
Month-to-month	2220	1655	3875
One Year, Two Year	2954	214	3168
Total	5174	1879	7043

An intervention can be planned to improve the recovery rate by increasing the duration of the customers' contract to one or two years. Considering that the Contract attribute is actionable, we can take it as the intervention T to make the customer increase the duration of the contract, Y as the outcome, meaning that the customer may or may not cancel the service and X as the set of features that may be interfering with the outcome.

3.2 Dataset preparation

To prepare the dataset, we loaded it into the R system and used a set of tools available in the Package 'Uplift' [14], as mentioned by Guelman *et al.* [15].

As shown in Table 1, only 1869 customers continue to use the services, which shows that churn distribution is unbalanced. In order to make the dataset more balanced, we use the **rvtu** (Response Variable Transform for Uplift Modeling) function. This function operates three main changes on the dataset: it creates a new response variable *y*, with a binary outcome corresponding to the one indicated to the function; it creates a binary variable *z* corresponding to the transformed outcome Y^* as indicated in the expression (9), and finally creates a new variable to assign each observation to the treatment (ct = 1) or control (ct = 0) group; in this case, the assignment was made by sampling without replacement, in order to distribute the observations among the treatment and control groups proportionally.

The result that relates the duration of the contract with the possible result (churn or not churn) after executing the **rvtu** function in the dataset is presented in Table 2.

Table 2. Number of customers that churn by Contract duration after **rvtu**

Contract attribute class	Churn attribute		Total
	Yes	No	
(0) Month-to-month	1828	1340	3168
(1) One Year or more	2954	214	3168
Total	4782	1554	6336

To carry out an exploratory analysis of the data and the importance of each of the attributes, we used the **niv** (net information value) function from Package 'Uplift'. This function finds the Net Information Value (NIV) and Net Weight Of Evidence (NWOE) for each of the attributes used in the model. Siegel [1] and Guelman [15] mentioned that the NIV and NOWE are Uplift measurements. The result obtained is shown in Fig. 4.

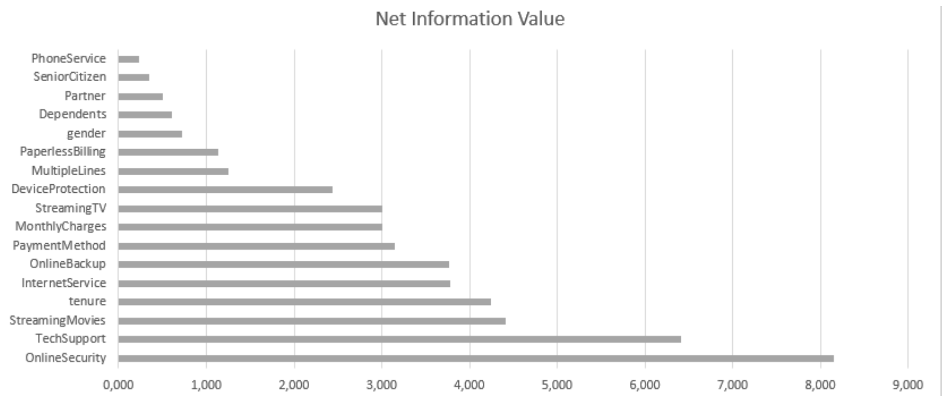

Fig. 4. Net information value by feature

3.3 Applied Model

Initially, and to later assess the quality of the model, the dataset was randomly divided into two parts, leaving 70% of the data for training the model and 30% for testing it.

To create the Uplift model, we use the **ctree** function of the Package 'partykit' [16]. As described in the documentation, **ctree**, short for conditional inference trees, is a binary recursive partitioning function for continuous, ordered, nominal, and multivariate response variables in a conditional inference framework. In this way, the model was built with the target variable Y^* which corresponds to the transformed result of expression (9). The tree obtained is shown in Fig. 5.

For each node and leaf of the tree, the CATE was calculated using the formula defined by expression (10). The values obtained are shown in Fig. 5, in the rectangles to the right and below each node and leaf, respectively.

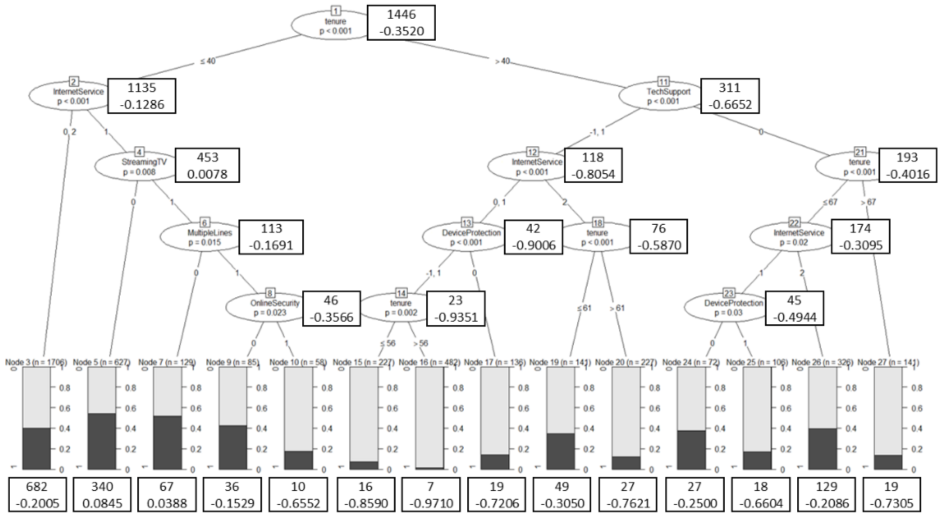

Fig. 5. The model obtained with the **ctree** algorithm for the Transformed Outcome approach

3.4 Evaluate Model

We used the Uplift chart to evaluate the developed model and calculated the MSE / PEHE as described in section 2.5.

To create the Uplift plot shown in Fig. 6, we score each of the observations in the test dataset with the corresponding uplift value obtained by the model. Then we sort the values obtained in ascending order and group the observations into ten groups (deciles). We find the ratio of those treated (ct = 1) for each decile and maintained the service (y = 1) concerning the number of customers in the decile. Likewise, we found the rate of those who were not treated (ct = 0) and maintained the service (y = 1) about the total number of customers in the decile. The difference between these two rates corresponds to the decile Uplift.

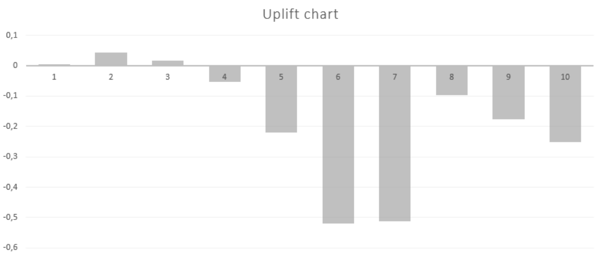

Fig. 6. Uplift chart of the Uplift model

To calculate the MSE, we consider the test dataset's Y* and the Uplift value found and described in section 3.3, applied in expression (12). The result obtained corresponds to MSE = 0.67.

As a general appreciation of the model, we found that although the Uplift graph does not present itself in the configuration of an ideal Uplift model, it presents a regularity and a descending curve between the 2nd and the 7th decile that suggests a suitable model. On the other hand, the value obtained for MSE does not correspond to the values of good predictive models. However, the value of MSE is in the same order of magnitude of the values of another works, like the ones in Zhang *et al.* [10].

4 Discussion: Persuadable profiles

Each branch of the decision tree shown in Fig. 5 forms a division rule, where each node includes one of the attributes used by the algorithm, seeking to optimize the prediction of the target Y*. Each node has a CATE value calculated as mentioned in section 3.3. Higher CATE values represent a more excellent cause-effect relationship between the intervention and obtained outcome. Therefore, it is expected that we will find the persuadable customers we are looking for in the leaves of the tree with the highest CATE. On the other hand, if we analyze the features and classes that give rise to the splitting of the tree's leaves that present a higher CATE, we can trace the profile of the most persuadable customers. The target of interventions should then be customers who present persuadable profiles above a CATE threshold. The available budget for the intervention should condition the limit above.

In this case study, we define CATE>0 as a limit. Profiles that have CATE>0 are shown in Table 3. The table also shows the number of customers for each node, those with month-to-month contracts, and the number of those who canceled the service. The last column considers the average monthly value of contracts of 61.46 dollars to present the maximum value that can be recovered if customers who have monthly contracts and have suspended the service respond positively to the intervention.

Table 3. Persuadable profiles

Node #	Customer profile to target	CATE		Customers #	Customers # Drop outs	Customers % Drop out	Max. recover value
5	Tenure<=40 and Internet Service=DSL and StreamingTV=No	0.0845	All	1041	307	29.49	
			M-to-m	846	296	34.99	18182.16
7	Tenure<=40 and Internet Service=DSL and StreamingTV=Yes and Multiple Lines=No	0.0388	All	180	37	20.56	
			M-to-m	115	27	23.48	1659.42

Given the above, an intervention aimed at customers who have a contract duration of fewer than 40 months, use the DSL internet service without Streaming TV or with Streaming TV but do not use multiple telephone lines in order to switch to annual or biannual contracts can lead to the recovery of 17.67% of dropouts (296 + 27 dropouts in profiles to target dividing by 1828 total dropouts with month-to-month contracts), corresponding to an increase in income up to 19,841.58 dollars per month. This value also represents a maximum estimate for the amount spent on recovering these customers without incurring losses.

5 Conclusions and Future work

Churn and how to deal with it is a big issue in the telecommunications sector. It has long been an area of research in predictive modeling, trying to predict the greater or lesser probability of a customer becoming a dropout. However, we think that this is just one aspect that can be considered. Within the scope of actionable knowledge, we argue that it is crucial to find effective and personalized interventions that can reduce dropouts and, at the same time, allow the causal effect of these interventions to be determined.

Actionable attributes are attributes that can be manipulated and allow operational changes. Based on the available attributes of the Telco dataset, we consider the contract type as an actionable attribute, as customers with monthly contracts tend to be less loyal, unlike those with longer-term contracts (annual or biannual).

Considering this personalized intervention that encourages clients to opt for a long-term contract in exchange for benefits, we use Uplift modeling as a machine learning-based technique to find actionable profiles and to determine the effects of treatment at the individual level in order to increase retention and find the right balance between the budget allocated to the campaign and the result obtained. The contribution of this work is therefore to find persuasive customer profiles, meaning customers who respond if and only if they are subject to the campaign (treated), through the application of uplift methods in order to obtain actionable knowledge.

However, in order to find a more substantial contribution to the study of Uplift models, it will be useful to carry out further studies to improve the results obtained. More work is needed at the dataset level, selecting subsets of features and pre-processing the data to obtain classes that have higher Uplift values and thus obtain trees that can present more useful actionable profiles. It may also be useful to investigate and experiment with other methods described in the Uplift literature, as well as other ways of evaluating the results obtained, and also comparing results obtained in other similar studies in the area of telecommunications, if available.

References

[1] E. Siegel, *PREDICTIVE ANALYTICS The power to predict who will Click, Buy, Lie, or Die*, Revised an. Wiley, 2016.

[2] E. Ascarza, "Retention futility: Targeting high-risk customers might be ineffective," *J. Mark. Res.*, vol. 55, no. 1, pp. 80–98, 2018.

[3] J. Pearl, M. Glymour, and N. P. Jewell, *Causal Inference in Statistics: A Primer*, 1st ed. Wiley, 2016.

[4] G. W. Imbens and D. B. R. Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences: an introduction*. Cambridge, 2015.

[5] J. Pearl and D. Mackenzie, *The book of why: the new science of cause and effect*. Basic Books, 2018.

[6] M. A. Hernán, J. Hsu, and B. Healy, "A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks," *Chance*, vol. 32, no. 1, pp. 42–49, 2019.

[7] F. Devriendt, D. Moldovan, and W. 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*, vol. 6, no. 1, pp. 13–41, 2018.

[8] P. Gutierrez and J.-Y. Gérardy, "Causal Inference and Uplift Modeling: A Review of the Literature," *Jmlr*, vol. 67, pp. 1–13, 2016.

[9] P. R. Rosenbaum and D. B. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, vol. 70, no. 1, pp. 41–55, 1983.

[10] W. Zhang, J. Li, and L. Liu, "A Unified Survey of Treatment Effect Heterogeneity Modelling and Uplift Modelling," *ACM Comput. Surv.*, vol. 54, no. 8, pp. 1–36, 2022.

[11] S. Athey and G. W. Imbens, "Machine Learning Methods for Estimating Heterogeneous Causal Effects," no. April, 2015.

[12] M. Jaskowski and S. Jaroszewicz, "Uplift modeling for clinical trial data," in *ICML 2012 Workshop on Clinical Data Analysis*, 2012.

[13] BlastChar, “Telco Customer Churn,” 2018. [Online]. Available: https://www.kaggle.com/blastchar/telco-customer-churn. [Accessed: 01-Nov-2021].

[14] L. Guelman, "Package 'uplift,'" 2015.

[15] L. Guelman, M. Guillén, and A. M. Pérez-Marín, "Optimal personalized treatment rules for marketing interventions : A review of methods, a new proposal, and an insurance case study," *Research Group on Risk in Insurance and Finance*, 2014.

[16] T. Hothorn, H. Seibold, and A. Zeileis, “Package ‘partykit,’” 2021.