Using behavioral analytics and machine learning to improve churn management

By

Ignacio Aberg Cobo

Bachelor of Science, Business Economics
UTDT, 1999
Executive Master in Business Administration
IE, 2008

SUBMITTED TO THE MIT SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN MANAGEMENT OF TECHNOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2017
©2017 Ignacio Aberg Cobo. All rights reserved.

The author hereby grants to MIT permission to reproduce
and to distribute publicly paper and electronic
copies of this thesis document in whole or in part
in any medium now known or hereafter created.

Signature of Author: __________________________

MIT Sloan School of Management
May 12, 2017

Certified by: __________________________
Duncan Simester
NTU Professor of Marketing
Thesis Supervisor

Accepted by: __________________________
Johanna Hising DiFabio
Director
Sloan Fellows and EMBA Programs
MIT Sloan School of Management
Using behavioral analytics and machine learning to improve churn management

By

Ignacio Aberg Cobo

Submitted to MIT Sloan School of Management on May 12, 2017 in Partial Fulfillment of the requirements for the Degree of Master of Science in Management of Technology.

ABSTRACT

New trends are shaping the telecommunications, media and technology (TMT) industries. Consumers are demanding to be connected anytime to hundreds of thousands of applications that are one click away. In addition, loyalty levels are decreasing and customers do not hesitate to switch providers if they do not receive value for their money. Because of this, churn management is a key driver of profits. However, few companies excel at churn management and most underestimate its impact.

The thesis is focused on describing a technological solution targeted to improve churn management capabilities within companies that belong to the TMT sector and explore the opportunities and hurdles of selling this kind of solution in a B2B context.

The hypothesis is that a world class churn management solution can effectively deploy statistical models to score customers by their likelihood to churn and execute targeted treatments for each segment through the operator service channels.

The study will focus on how behavioral analytics and machine learning can increase customer’s life time value and boost margins in TMT companies. Throughout the research, I will describe the best practices within the industry to establish a state of the art churn management solution.

Thesis Supervisor: Duncan Simester
Title: NTU Professor of Marketing
Table of contents:

1. Introduction 6
   1.1. Churn management opportunity 7

2. State of the art churn management solution 10
   2.1. Predictive analytics 11
   2.2. Discount optimization 16
   2.3. Online dashboard 21

3. Strategic framework to evaluate solution 24
   3.1. Create value 25
   3.2. Capture value 29
   3.3. Deliver Value 31

4. Final conclusion 34

5. Appendix
   5.1. References and bibliography 36
   5.2. Citations 36
Zig Ziglar, American Author: “Statistics suggest that when customers complain, business owners and managers ought to get excited about it. The complaining customer represents a huge opportunity for more business”.

1. Introduction

Nowadays, customers are more active than ever and are demanding value for their money and are willing to switch providers if they perceive the service does not meet with their expectations. In the telecommunications space, customers are demanding to be connected anytime to hundreds of thousands of applications that are one click away and more than ever before, loyalty levels are decreasing.

Because of these factors, churn management and retention discount optimization are two the most important drivers for profits maximization in the telecommunications, media and technology (TMT) industry.

The industry is facing a consolidation phase because growth is starting to stagnate and firms are having much more pressure on their bottom line. As a consequence, organizations need to adopt analytics capabilities because if not they are at risk of not surviving as the whole industry evolves towards data driven solutions.

Before joining MIT, my role at my former employer involved reducing the churn of the customer base. I had to ensure retention success of over 240,000 monthly cancellation requests that were handled by 1,200 CSRs across Latin America. Additionally, I monitored the retention discounts and credits granted to clients that represented almost US$100M annual credits.

As part of that experience, I recognized the huge opportunity to improve the retention process of subscription-based services in Latin America and while studying at MIT I have broadened my perspective about analytics and machine learning.

My professional and academic reflections are presented in this paper and the aim of the research is to describe how the retention process can be improved.

“Churn” is the counterpart of retention. If the customer has decided to stop transacting with the firm, the customer has churned. In that sense churn is inferred by the cessation the customers’ transactions with the firm.(Ascarza et al., 2017)

The thesis is focused on describing a technological solution targeted to improve churn management capabilities within companies that belong to the TMT sector and explore the opportunities and hurdles of selling this kind of solution in a B2B context.

The paper is divided in two main sections. In the first two chapters the goal is to understand the opportunity around marketing analytics tools and analyze how churn management practices can positively impact customer lifetime value and increase margins in TMT companies.
Furthermore, this section aims to answer the question that every marketing manager should be asking in the industry, that is the following: “Which is the statistical model that allows me to extract the most value from my existing customer data?”

The answer to this question is related to algorithms that help marketing managers retain their existing customers and optimize the amount of discounts each of these customers should receive when calling to cancel the service.

In order to support these ideas, the paper describes a churn management state of the art solution and how world class churn management companies can effectively deploy statistical models to score their customers by their likelihood to churn and execute targeted treatments for each segment by applying analytics and machine learning techniques and institutionalizing a process of randomized experimentation to find the causality of the key variables that optimize the retention discounts while maintaining a sustainable save rate.

In the second part of the paper, the goal is to describe the obstacles to deploying this churn management solution to companies within the TMT sector in Latin America. The argument assumes that there is a huge opportunity to offer this solution to Tier 2 and Tier 3 companies that are willing to invest in technology to improve their margins.

This section analyzes the solution with the lens of a strategic framework that is used at MIT Sloan School of Management to evaluate business opportunities. This framework is based on the create, capture and deliver value structure and helps to achieve a consistent strategic thinking.

Finally, the paper concludes with a general overview of the industry and an urgent call to TMT executives requesting them to take risks and adopt data driven technologies in order to create sustainable business models that enhance the customer’s experience.

1.1 Churn Management opportunity

Currently, there is phenomenal interest in artificial intelligence and machine learning projects. In the United States, business media and specialized reporters are constantly writing articles about companies from the Fortune 500 index that are investing heavily in data capabilities and infrastructure to gain a competitive advantage by generating insights from the information available.

In addition, many start-ups have been funded during the last couple of years to provide analytics as a service to major companies who have not yet figured how to use the data. Figure 1 presents 100 data driven start-ups that are trying to transform several industries by applying artificial intelligence and machine learning.

The figure illustrates how the new ventures are trying to differentiate themselves by specializing in various industries and particular capabilities or functions in the value chain.
The main reasons why so many companies are focusing on data science are related to the increase in computational power and the amount of data digitalized during the last few years. These two major trends are allowing companies to apply in business settings the statistical methods discovered decades ago, like for example, logistic regression, k-nearest neighbor, logic trees and neural networks.

In this context, there is a huge timing opportunity to develop a solution targeted to the postpaid subscriber in the TMT sector by using analytics and experimentation to reduce churn and optimize retention treatments.

The Perils of proactive churn prevention using plan recommendations paper states that “The annual churn rate for wireless telecoms providers is approximately 15%-30% worldwide, which has been estimated to cost organizations up to $10 billion annually. As a result, companies are increasingly managing customer retention proactively by identifying valuable customers who are likely to churn and taking appropriate action to retain them” (Ascarza, Iyengar, & Schleicher, 2016)

In Latin America, TMT companies need to better analyze the trade-offs between customer acquisition and retention in order to improve their margins. Usually companies spend more resources on newly acquired customers than on retained customers and as a consequence, when the selling promotion ends, these newly acquired customers tend to call the service providers to re-new their discount.

A famous business maxim states that “it costs less to retain an existing customer than to find a new one” but nevertheless, churn ratios are high in the industry and I estimated from interviews and my professional experience that telecommunications and cable companies spend on average between
5%-10% of their revenues on retention discounts. As a result, discount offerings have a huge impact on the bottom line of these companies.

The following table represents an illustrative example of three hypothetical operators with different company size and average industry indicators, like for example, ARPU, cancelation requests, save rate and percentage of discounts granted per year.

<table>
<thead>
<tr>
<th>Illustrative example of three players:</th>
<th>Operator 1</th>
<th>Operator 2</th>
<th>Operator 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARPU ($)</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Average Post paid (Subscribers in thousands)</td>
<td>10,000</td>
<td>5,000</td>
<td>1,000</td>
</tr>
<tr>
<td>Post paid revenues per year ($ in millions)</td>
<td>9,000</td>
<td>3,000</td>
<td>6,000</td>
</tr>
<tr>
<td>Cancelation requests per month (%)</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Cancelation requests per year (Subscribers in thousands)</td>
<td>3,600</td>
<td>1,200</td>
<td>240</td>
</tr>
<tr>
<td>Save rate (%)</td>
<td>80%</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Customers retained per year (Subscribers in thousands)</td>
<td>2,880</td>
<td>960</td>
<td>192</td>
</tr>
<tr>
<td>Customers retained with retention discounts per year (%)</td>
<td>80%</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Customers retained with retention discounts per year (Subscribers in thousands)</td>
<td>2,304</td>
<td>768</td>
<td>154</td>
</tr>
<tr>
<td>Retention discounts granted per year over revenues (%)</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Retention discounts granted per year ($ in millions)</td>
<td>450</td>
<td>150</td>
<td>30</td>
</tr>
<tr>
<td>Reduction Retention discounts granted per year (%)</td>
<td>15%</td>
<td>15%</td>
<td>15%</td>
</tr>
<tr>
<td>Reductioin Retention discounts granted per year ($ in millions)</td>
<td>67.5</td>
<td>22.5</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Source: Interviews, personal experience

For example, in the case of an operator with 10 million subscribers we estimate that on average there are 3,600,000 cancellation requests per year. But the most relevant factor is that 2,304,000 of those customers who called to cancel the service, received retention discounts representing at least 5% of the revenues of the company.

A solution that can optimize retention discounts granted without reducing the save rate can generate important costs savings for TMT companies. For instance, assuming a conservative ratio of 5% of discounts over revenues, a reduction of 15% represents savings of $67.5M in the case of 10 million subscriber base operator (Operator 1 in the table above). If we consider a smaller operator of 1 million users and follow the same rationale described above, we can obtain cost savings of $4.5MM a year (Operator 3 in the table above).

As a conclusion, we observe that a reduction of 15% in the retention discounts budget generates savings of $67.5M, $22.5M and $4.5M for operators 1, 2 and 3 in the example of table 1.

The main take away of this section is that there is a strong business case that supports the idea of providing a retention discount optimization solution to TMT companies and that these companies
should be willing to invest to better handle cancelation requests and minimize their retention costs without reducing the save rate ratio.

2. State-of-the-art churn management solution

The main characteristics that describe the churn optimization challenge are the following:

- **Timely:** When a customer calls to cancel a subscription, the service provider needs to decide whether or not to assign a retention offer to that specific customer
- **Data-driven:** The service provider knows relevant information about the customer because it is an existing customer who has a relationship with the company
- **Meaningful:** The churn management process has a huge impact on the bottom-line of the company

This context generates a unique opportunity to develop a solution with analytics that can help companies make better decisions when treating this kind of customer.

An Accenture’s publication of 2011 about maximizing customer retention describes a strategic approach to effective churn management that requires a new set of capabilities for TMT companies: “Acquire a detailed, fact-based understanding of customers’ intentions and what makes them switch; get offers to market swiftly, using a rapid “test, learn, and scale” mode; deploy real-time treatment tools across customer interaction channels that will ensure the right retention decisions by weighing customer churn propensities against customer value. Accenture’s global research exposes the scale of the challenge. Despite a small overall decline in switching levels, two in three customers have changed providers in the past year in at least one of the industries covered in the research because of dissatisfaction with service levels”. (Hanson, 2011)

The state-of-the-art solution proposed in this paper is consistent with Accenture’s concept of world-class churn management and consists of the following three components:

- A predictive algorithm to predict who is at risk, i.e. who are the customer more likely to cancel the service?
- A discount optimization solver to help the marketing managers decide how much to invest in these customers
- An online dashboard to monitor and control the evolution of discounts granted and the evolution of the customers who churned

The following figure illustrates the three components of the solution that allows companies to answer these questions: (i) which customer is at risk, (ii) how much to invest in each customer and (iii) how to control and monitor the treatments applied to customers.

The sustainable advantage of the solution is that the accuracy improves over time as the machine learns with more data which means that the solution is impossible to be reversed engineer by another company.
Moreover, the application is embedded in the decision making process in order to add a measurable value to the TMT service provider.

The following section defines in more detail the three components of the solution and describes how each one works and interacts with the other components.

2.1 Predictive analytics

The predictive analytics component consists of a statistical model that predicts who is going to churn and estimates the probability of churn for each customer. The model analyzes internal information about the customer, for example:
- usage of the product
- transactional and demographics variables
- method of payment
- contacts with the company by reason (technical, billing, collections, etc.)

In case adding external information about the customer is possible, this could improve the accuracy of the prediction.

The outcome of the algorithm is a propensity or probability from 0 to 1 for each customer to call to cancel the service. Additionally, the algorithm will cluster the customer in different segments, for example, low engagement with the product, discount credits addicts, or high margin customers.
This solution will help service providers to identify the drivers of churn and to accurately correlate churn to descriptive variables.

The following figures help to explain in a visual way how these functions work.

**Figure 3: Predictive analytics solution**

*Methodology to predict customers likely to leave in next 30 Days*

![Diagram](image)

Source: Past professional experience and former employer presentations; google images

In order to facilitate the implementation of a solution of this kind, the suggestion is to use a cloud storage service to gather and store the information because it is easy to implement, execute and scale. As stated in The Economist magazine: "Thanks to Amazon Web services there has never been a better time to start up a web based or data centric-firm".
The churn management solution has three components and makes commercial sense to target the first component of the solution to tier 1 TMT companies because almost every company of this group possesses a CRM with real time data of their customers meaning that the gathering of the information is not an obstacle to develop this kind of model. But this component is very difficult to sell to this cluster of TMT companies because almost every tier 1 operator has already developed churn predictive models that gather data from internal sources.

In contrast, there is a huge opportunity for offering this predictive component to tier 2 and tier 3 companies in Latin America because these companies do not have the skills and capabilities to develop this kind of predictive model. Furthermore, these organizations do not have the financial resources to acquire a solution from top technological enterprises, like for example, SAS, IBM or Teradata.

To make the solution scalable, the offer should consist of a presetting of algorithms that can be configured according to the customer request or focus. The key to success is developing a portfolio of algorithm and replicate the McDonald's menu concept. As a consequence, the companies buying the predictive component of the solution can select which kind of algorithm they want to utilize on their customers' base.

The predictive algorithms offering should be 80% standardized and 20% customized by TMT operators and possess three different outcomes:

- Propensity score for each customer to churn
- Forecasting variables, like cancellation requests, churners or discounts granted
- Segmentation of customers
Figure 5 presents a visual explanation with examples on how to develop a menu of standard algorithms related to different churn management solutions.

**Figure 5: Standardized algorithms by outcome**

<table>
<thead>
<tr>
<th>McDonald’s Menu of Combos</th>
<th>Set of standard algorithms related to churn</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>extra value meals</strong></td>
<td>• Propensity score algorithms</td>
</tr>
<tr>
<td></td>
<td>• Techniques: Linear regression, neural nets, k-nearest neighbor</td>
</tr>
<tr>
<td></td>
<td>• Outcome: Risk score for each customer</td>
</tr>
<tr>
<td><strong>Big Mac</strong></td>
<td>• Forecasting algorithms</td>
</tr>
<tr>
<td><strong>Any Size Fountain</strong></td>
<td>• Techniques: Linear regression, neural nets, k-nearest neighbor</td>
</tr>
<tr>
<td></td>
<td>• Outcome: Cancelation requests prediction</td>
</tr>
<tr>
<td><strong>Juicy Lucy</strong></td>
<td>• Segmentation algorithms</td>
</tr>
<tr>
<td><strong>Grilled Chicken Tender</strong></td>
<td>• Techniques: Neural nets, classification trees, Logistic regression, Naive Bayes, Discriminant analysis</td>
</tr>
<tr>
<td><strong>Non-Fat Chicken Tender</strong></td>
<td>• Outcome: Clusters of customers</td>
</tr>
</tbody>
</table>

The following figure presents illustrative examples of different types of segments. The importance of clustering customers by risk, value and behavior is related to assigned specific treatments to each group so as to maximize their respective customer life time value.
In summary, with the statistical model we will find:

- Who will churn in the near future? (prediction)
- What is the likelihood of churn of each customer? (Score)
- How much is the actual and potential value of the customer? (monetary value)
- To which segment each customer belongs to? (clustering)

TMT companies are urged to develop a world class churn management solutions based on data-driven capabilities and this statement is aligned with the recommendations of the top consultancy firms, for example PricewaterhouseCoopers' published an article about curing customer churn in 2011 that details the benefits of using analytics to improve churn: “Using efficient statistical methods, we combine data collected from diverse systems, call monitoring, and primary research surveys with elements from a customer data-mart (account, usage, and revenue data). The customer data mart can be an existing data store or an ad hoc database created quickly to support analysis. Data analytics is most powerful and accurate when combined with the other three tools—process analysis, call monitoring, and primary research—because these distinguish statistical correlations from true cause/effect relationships. This creates an accurate timeline of those events in the customer life cycle that ultimately lead to churn”.(Abbott & Karakaya, 2011)

The paper “In Pursuit of Enhanced Customer Retention Management: Review, Key Issues, and Future Directions” (Ascarza et al., 2017) presents more advanced predictive algorithms using machine learning and big data modeling. These developments could play key roles in developing predictive models for who is at risk, why, who will respond, and when to target. Deep learning, for example, is a machine learning approach based on neural networks that combines supervised and unsupervised aspects. Deep learning has been used to learn about customer probability to defect (Castanedo et al. 2014), and may also be helpful in modeling response to retention offers. Similarly,
boosted varying-coefficient regression models have been studied for dynamic predictions and optimization in real time (Wang and Hastie 2014) and have been shown to offer a major improvement over the classic stochastic gradient boosting algorithm already used for churn prediction (Lemmens and Croux 2006; Lemmens and Gupta 2013).

2.2 Discount optimization

The second component of the solution is a retention discount optimization tool that is the most compelling device of the solution.

It is relevant to recall that almost every telecommunication and cable company spends on average 5% of their revenues on retention discounts, which presents a huge business opportunity to optimize the discount offering and impact the bottom line of these companies.

So how does it work? Figure 7 illustrates a situation where a customer calls to cancel the service and the call is handled by a customer service representative. This is the moment of truth for the relationship between the customer and the service provider, so the main idea is to develop a solution that can help the customer service representative make better decisions about how to handle that call.

![Figure 7: Discount optimization: how does it work?](image)

The solution will provide real-time discount recommendations that help the CSRs to manage the customer's cancelation requests and to achieve this, the software needs to determine the optimal retention offer for each interaction for every customer.
The following figure illustrates the different types of discounts that telecommunication and cable companies are granting. As we can observe, there are multiple combinations of discounts that present an optimization challenge to the service providers.

**Figure 8: Optimization challenge**

Types of discounts that can be customized accordingly to clients request

<table>
<thead>
<tr>
<th>Monthly Subscription</th>
<th>Description: Discounts associated to the monthly subscription expressed in % of discount and duration of the offer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Examples: 15% x 3 months, 15% x 6 months, 25% x 3 months</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technology / Equipment</th>
<th>Description: Discounts granted to help customers upgrade their equipment or technology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Examples: DVR HD free, Samsung s6 at 50% off</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Premium or value added services</th>
<th>Description: Discounts linked to premium services that enhance the customer experience</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Examples: 100% HBO, 50% FOX Sports, 50% cellphone premium content</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>“One shot benefit”</th>
<th>Description: One shot credit notes to fix a particular issue that a customer has faced with the company</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Examples: Free 100 minutes, $15 credit note</td>
</tr>
</tbody>
</table>

Source: Past professional experience and interviews

The proposal to solve the discount optimization challenge is the most compelling insight of the research because it is based on the iteration of the following two capabilities:

- **Experimentation**: to understand how customers are responding to the retention treatments. This analysis takes into account the clustering and segmentation that we defined in the first component and the logic is to use the responses of the treatments as an input to the optimization algorithm
- **Optimization algorithm**: based on optimization techniques, like for example, linear and integer programming algorithms

The benefits of running controlled experiments are the opportunity to test different hypotheses about the customer’s preferences and to obtain continuous feedback on the reactions of the customers to different treatments. In order to promote a culture of experimentation, it is necessary to conduct thousands of experiments that help to improve how a company should retain its customers.

Moreover, the proposal implies that the analytics team should embrace the test-and-learn approach in order to get feedback from the experiments and as a consequence, obtain improvements in profits by reducing the discounts granted to customers.

The following figure presents a description of the iterative process that links the experimentation with analytics.
The competitive advantage of the iterative process is that the algorithm continuously learns and improves as it gathers more information about the customers and incorporates the outcomes of the experiments. This process enhances the institutionalization of a culture of experimentation within the company in order to improve the accuracy of the algorithm.

The main take-away of this section is that the iteration between experimentation and analytics is key to identify the causality of the treatments that have a positive correlation with the objective function we are trying to solve with the algorithm (minimize retention discounts).

The following figure presents an example of a simple A/B test to determine price elasticity to a specific retention discount.
Figure 10: Institutionalize experimentation

A/B Testing Illustrative example

**Treatment group:** Customer A, churn score 80 → Cancelation request → Option A: Retention Offer recommended: 20% discount for 3 months

**Control group:** Customer B, churn score 80 → Cancelation request → Option B: Retention Offer recommended: 10% discount for 3 months

Source: Past professional experience and interviews

Figure 11 provides an example of a set of variables utilized to optimize an objective function.

**Figure 11: Optimization algorithm to reduce retention discounts**

<table>
<thead>
<tr>
<th>Examples of variables for the optimization solver</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objective function</strong></td>
</tr>
<tr>
<td>- Maximize each customer lifetime value by granting discounts that increase the retention rate</td>
</tr>
<tr>
<td><strong>Decision Variables</strong></td>
</tr>
<tr>
<td>- Binary Variables: decide whether to assign or not a retention offer to a customer who has presented a cancellation request</td>
</tr>
<tr>
<td>- # offers assigned to each customer</td>
</tr>
<tr>
<td>- Duration of retention offers (months)</td>
</tr>
<tr>
<td>- Average ARPU of customers retained ($)</td>
</tr>
<tr>
<td>- Average Discount of customer's retained ($)</td>
</tr>
<tr>
<td>- Customers with retention offers / total customers (%)</td>
</tr>
<tr>
<td>- Customers retained with retention offers (#)</td>
</tr>
<tr>
<td>- % of Voluntary churn targets and KPIs</td>
</tr>
<tr>
<td>- Relative price vs competition offers</td>
</tr>
<tr>
<td>- Repeatative customers with cancellation requests</td>
</tr>
<tr>
<td><strong>Constraints</strong></td>
</tr>
<tr>
<td><strong>Optimize retention discounts spending by:</strong></td>
</tr>
<tr>
<td>- Reducing discounts in low churn risk customers</td>
</tr>
<tr>
<td>- Rebalancing discounts from low value to high value customers</td>
</tr>
<tr>
<td><strong>Discount optimization problem:</strong></td>
</tr>
<tr>
<td>- linear programming</td>
</tr>
<tr>
<td>- Integer programming</td>
</tr>
</tbody>
</table>

Source: Past professional experience, former job presentations and interviews with TMT executives

In this example, the objective function is to increase each customer lifetime value by optimizing the amount of discounts granted to customers while increasing the retention rate. The optimization model should solve this business problem for each customer by maximizing "the present value of the future..."
The following formula is based in the CLV definition used in the course marketing analytics at MIT Sloan School of Management by Dean Eckles.

\[
CLV = \sum_{t=1}^{T} \left( \frac{\text{revenues}_{it} - (Y_{ij} \times \text{discount}_{jt}) - \text{cost}_{it}}{(1 + \text{discount rate})^t} \times (\text{Retention rate}_{it} \times \text{propensity score}_{it}) - \frac{\text{Acquisition cost}_i}{(t-1)} \right)
\]

Each variable of the formula means:
- CLV is the customer lifetime value of customer i at time t
- Revenues is the inflow from customer i at time t
- Y is a binary variable: if 1 we offer customer i a retention discount j, if 0 we do not offer a retention discount
- Discount is the retention offer j assigned at time t
- Cost is the cost of servicing i at time t
- Retention rate for customer i at time t
- Propensity score to churn of customer i at time t
- Acquisition cost is the initial cost to acquire customer i
- Discount rate is to discount for future profits

The decision variables are the following: (i) to decide whether or not to assign a retention offer to a customer who has presented a cancellation request and (ii) which type of discount to offer.

The constraints are very relevant to the algorithm and can be configured by the service provider according to the company's priorities. Some examples of possible constraint variables are the following:

1. **Customers characteristics:**
   a) Average ARPU of customers retained ($)
   b) Average Discount of customer's retained ($)
   c) Customers life time value
   d) Repetitive customers with cancelation requests
   e) Customer segment
2. **Offers restrictions:**
   a) Duration of retention offers (months)
   b) Customers with retention offers / total customers (%)
   c) Customers retained with retention offers (#)
3. **Internal KPIs:**
   a) % of Voluntary churn targets
   b) Sales target
4. **External information:**
   a) Relative price vs competition offers
b) New promotions granted by the competition

The following figure presents the conceptual trade-offs of granting retention discounts to a specific customer because the number of discounts positively affects the retention rate but it has a negative effect on the customer lifetime value.

![Figure 12: Optimization model – conceptual trade-offs](image)

A positive characteristic of this component of the churn management solution (Discount optimization) is that every marketing manager of a TMT service provider can easily measure the success of the algorithm by monitoring the evolution of the amounts of discounts granted and the retention level of their customers.

2.3 Online dashboard

The third and last component of the churn management solution is the online dashboard. This dashboard has two main functions:

- Monitor and report
- Configure and execute

First, the capabilities related to monitoring and reporting are oriented to help the marketing managers analyze the accuracy of the predictive algorithms by comparing actual outcomes to the predictions (cancellation requests and customer’s propensity to churn).

The following matrix (named Confusion Matrix) helps to explain the meaning of the accuracy of the predictive algorithms. In this use case related to churn management, the matrix compares the churn
ratio's accuracy which is the difference between the actual cases in the sample vs the prediction of the model.

<table>
<thead>
<tr>
<th>Table 2: Confusion Matrix</th>
<th>Predicted by the algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-churn</td>
</tr>
<tr>
<td>Actual: sample with historical information</td>
<td>Non-Churn</td>
</tr>
<tr>
<td></td>
<td>Churn</td>
</tr>
</tbody>
</table>

Additionally, the dashboard allows the marketing manager to monitor the offer effectiveness in a daily roll over dashboard. The dashboard follows a similar logic as the Google AdWords dashboard where the campaign manager can monitor and manage the ROI of every marketing campaign. This means that the churn management dashboard helps to identify the money invested in every customer and estimate the customer lifetime value. This visualization characteristic is very useful to help the marketing managers understand the profitability of each customer during the different stages of the relationship with the TMT service providers.

This concept of analyzing the investment in each marketing campaign is associated with the approach to retention marketing stated in the article Retain Customers and Win How a Comprehensive Churn Management Approach Improves Margins by North Highland: "Retention marketing requires practitioners to advance their thinking from traditional marketing doctrines. The distinction is that companies must market, not only to expand market and wallet share, but also to hold onto their existing share. Companies that do this enjoy the benefits of customers that stick around longer and spend more money. Retention marketing is an extremely powerful tool that nets measurable results and returns significant value to organizations with the foresight to employ it". (Hawkins, n.d.)

The following figure contains an illustrative example of the "look and feel" of a dashboard:
The second function of the online dashboard is dedicated to configure and execute.

One of the main issues that marketing managers usually complain about regarding IT solutions is that each change in the marketing strategy requires a modification in the solution that takes time and does not take into account the urgency of a competitive market.

To avoid this common user complaint, the online dashboard needs to have the capacity and flexibility to allow changes and configurations of the strategy by the marketing manager without demanding any adjustment to the software.

To summarize this concept, the online dashboard (third component) would allow an ease to execute parametrization of the variables that affect the first and second component of the churn management solution.

The following list includes some examples of adjustments and configurations that a marketing manager could handle by using the third component of the solution:

- Predictive analytics:
  - Include new variables to the predictive models, like for example, external information or the outcome of customers surveys
  - Develop new segmentation rules according to business parameters, such as new cutoffs based on profitability or margins criteria
  - Configure the treatments to different customer segments
- Discount optimization algorithm:
o Adjust the decision variables of the optimization algorithm
o Define new business rules to limit or expand the retention discounts granted to different segments
o Create new retention offers as a defensive move against a competitor's aggressive marketing campaign
o Add new constraints to the optimization model that reflect a new scenario, such as budget limitations

Finally, the online dashboard is a key element to institutionalize the culture of experimentation within the operators because it helps the marketing managers to execute, measure and understand the results of the experiments in a consistent and structured way.

The goal of this section was to present the opportunity around data-driven tools and describe a world class churn management solution for TMT companies. The main take away is that a solution of this kind can positively impact customer lifetime value and increase margins in TMT companies.

The next section examines the opportunities and hurdles of selling this kind of solution in a B2B context by applying a high level strategic framework to assess the competitive advantage of the solution.

3. Strategic framework to evaluate the churn management solution in Latin America

This chapter analyzes the churn management solution with the lens of the strategic framework that is used at MIT Sloan School of Management to evaluate business opportunities. This framework is based on the "Create, Capture and Deliver Value" structure that is described in the following figure.

![Figure 14: Strategic framework: Create, capture and deliver value](image_url)

Source: MIT Sloan School of Management, Duncan Snider: course 15.809 Marketing and Strategy
For the purpose of this thesis, I will assume that the solution is developed by a hypothetical start-up named “Newton” and I will apply the three components of the strategic framework to analyze the challenges and obstacles related to deploying this churn management solution in companies within the TMT sector in Latin America.

3.1 Create value

Creating value is an approach that focuses on understanding how much value can be created by Newton and evaluating if TMT operators would be willing to acquire Newton’s technology in order to improve their margins. The goals of this section are the following:

1. Analyze the total addressable market for this solution within the TMT sector
2. Examine the purchasing decision process of the managers who decide whether or not to buy a solution of this kind
3. Assess if TMT operators will recognize the value created by the solution

The first step in this section is to study the total addressable market for this solution within the TMT sector. The potential buyer of this solution is almost every service provider company that needs to improve its churn predictive model and is capable of gathering data from internal sources. The following figure illustrates the potential users of the solution grouped in three clusters according to the market size of the companies.

![Figure 15: Total addressable market in Latin America by tier](image)

For the purpose of this thesis, the segmentation of the market follows the resulting rationale:
• Tier 1 organizations: Operators with more than 10 million customers located in populated countries with > 40 million inhabitants, like for example: Argentina, Brazil, Colombia and Mexico

• Tier 2 organizations: Operators with more than 1 million customers located mainly in countries with population > 15 million inhabitants, like for example: Chile, Ecuador, Peru and Venezuela

• Tier 3 organizations: Operators with more than 100,000 customers located in less populated countries with population < 5 million inhabitants, like for example: Caribbean islands, Puerto Rico and Uruguay and regional players in populated countries

These TMT operators represent a huge market because they provide multiple services and bundles of products, such as landline, mobile, broadband and pay TV offerings.

The second step of this section is to understand the purchasing decision process of TMT operators for a solution of this kind and analyze the differences between the three different tiers.

In tier 1 organizations, many people are involved in the decision to purchase a churn management solution and the purchasing process is time consuming. The following figure exemplifies the managers involved in this decision within tier 1 organization. The illustration identifies three different groups with different interests and backgrounds. For example,

• Data scientists of the analytics team are worried about: *How much time to process the data and get the insights?*

• Marketing managers are interested in achieving business results and they need to know: *How much time do they need to obtain the business results?*

• IT managers are concerned about: *How many resources should they allocate to set up and deploy this solution? Is the data clean and available?*
Newton as an early start-up with a new solution should not target tier 1 operators because of their complex decision process. Instead, Newton should target tier 2 and tier 3 operators because their decision process is much simpler and quicker which makes this segment much more attractive to approach.

This strategy may seem counterintuitive since almost every tier 1 firm possesses a CRM with real time data of their customers which means that the data gathering is in place but a start-up offering this solution should follow a two-stage commercial strategy by targeting tier 2 and tier 3 operators in the first phase and offering the solution to tier 1 operators in the second phase. Figure 17 describes the two stages of the commercial strategy.
The third step is to analyze if TMT operators will recognize the value created by the solution developed by Newton, so the question we should be asking is: "How can Newton capture the attention of TMT operators' managers and make them choose Newton's solution among other options."

When searching for the best alternative to buy, almost every corporate manager requests the vendor to provide a proof of concept of the solution before deciding which solution to choose, so developing a proof of concept is key to generate the buy-in from the operator.

The proof of concept is a demo of the solution conducted with a small sample of the operator's data that includes a business case with the economic benefits obtained by implementing the solution and the rationale regarding how the benefits are captured.

In order to generate a business case, it is necessary to involve the operator's marketing and analytics teams to gather and structure the data. According to the interviews I conducted during this research, data scientists estimate that a business case supporting the proof of concept could be generated in two weeks' time if the data has already been gathered and structured.

The ability to communicate and develop a business case to support the benefits of the churn management solution is a fundamental skill that a start-up needs to develop.

Once TMT operators' managers validate the demo of the solution and accept the economic benefits presented in the business case, the next step of the selling pitch should be focused on the concept of "paying for performance." This means that the corporate customers will pay only for results related to...
the discounts optimization, for example, for every dollar generated in savings, there would be a 25% fee for the solution provider, Newton.

Because of this, it is key to agree to the terms to evaluate the success of the solution before the implementation, and the comparison of the savings obtained with the solution should be against a randomized control group that is not affected by the optimization algorithm.

The practices mentioned above (i) proof of concept and (ii) pay for performance are determinant to the success of a start-up that is offering this kind of B2B solution to TMT service providers.

It is important to mention that tier 2 and tier 3 companies have not developed the capabilities required to search and compare different IT solutions, and because of this, it is much easier to be compatible with their buying process.

### 3.2 Capture value

Capturing value is an approach that focuses on horizontal competition and this section describes the main competitors for this solution and how this solution can sustain a differentiation advantage for Newton. (Simester, 2016a)

Regarding the competition, many well-established companies are providing predictive analytics products and real-time discount recommendation solutions to tier 1 operators with added value services such as consulting and 24x7 support. These technological leaders are not aiming their products to tier 2 and tier 3 companies because they do not consider this market to be profitable to support their large operating costs. Because of the fierce competition on the tier 1 segment, there is a timing opportunity for a start-up to serve the tier 2 and 3 segment in Latin America before other players decide to enter this niche.

Newton, as an early start-up with a new solution, should capture this opportunity by targeting tier 2 and tier 3 operators in Latin America who cannot afford to acquire technology from well-established IT players. This approach is aligned with the two phase commercial strategy mentioned in the previous section.

The following figure illustrates some examples of technological companies that are competing in the tier 1 market.
The potential competitors to Newton are other start-ups with great proprietary products that might be interested in the Latin America market and are willing to offer their solution at a reasonable cost that can be afforded by tier 2 and tier 3 players in Latin America.

So how can the solution developed by Newton differentiate from other start-ups or products targeted to tier 2 and tier 3 segment in Latin America? And how can this advantage be maintained over time?

To generate a sustainable business model with a competitive advantage, Newton needs to develop the following two strategic resources (i) relationships and (ii) switching costs:

- **Relationships with TMT operators**: generating trust as a strategic resource by building on positive results of initial sales and product robustness ("relationship resource")

- **Switching costs for the TMT operator by:**
  - offering high quality consulting services to deploy the solution
  - deploying machine learning algorithms that increase the speed to obtain business results and cannot be reversed engineered by another competitor

If Newton is capable of generating these two strategic resources, Newton will be the owner of the customer. In this case because it is a B2B solution, the customers are the TMT operators.

It is important to mention that both creating and capturing value are consistent with the hypothesis that selling to tier 2 and tier 3 companies is more efficient at the early stage of a start-up.

3.3 Deliver value
Delivering value is an approach that focuses on collaborating and competing in the value chain. (Simester, 2016b). In this section, the objective is to explore two kinds of partnerships, first opportunities around the supply side and secondly, opportunities around the demand side.

The hypothesis regarding the supply side is that a start-up will need to gain technical capabilities to develop the solution and the fastest way to learn these skills is to partner with a specialized company in machine learning that is not competing in Latin America.

Regarding the demand side, the hypothesis is related to develop a go-to market strategy in two phases that does not require a partnership in the first stage but the second phase may involve working with partners such as consulting firms. Figure 19 presents the structure and hypothesis of this section.

From the supply side perspective, the hypothesis is that data-driven solutions with machine learning techniques require technical expertise that a start-up needs to acquire, so Newton as a start-up that aspires to develop this solution, needs to partner with another player in order to gain the technical capabilities required for the challenge.

A partnership with a company that is focused on machine learning will accelerate Newton’s learning curve on technology and will help Newton to capture the timing opportunity in Latin America. Because of this, it is key to partner with a company that has no knowledge of the Latin American market so that Newton can contribute that expertise.

An example of a potential partner is DataRobot, a Boston based tech company that offers a machine learning platform for data scientists of all skill levels to build and deploy accurate predictive models in
a fraction of the time as stated in the company webpage. According to Venture Wire, DataRobot has raised another round of funding. The latest round of funding is a $21 million Series A that values the company at more than $60 million.

DataRobot has developed cutting edge technology because it is located in the analytics hub near MIT where data scientists and computer scientist are highly qualified and easy to find but this company has not developed a commercial priority towards Latin America. This situation represents an opportunity for Newton to partner with DataRobot and offer data driven marketing solutions to tier 2 and tier 3 companies in Latin America.

The following figure illustrates the main advantages of Data Robot’s predictive solution.

As we mentioned in the previous sections, Newton’s commercial strategy should have two phases: during the first phase, Newton should target tier 2 and tier 3 companies because the selling process is simpler than tier 1 companies and after being successful in this niche, Newton should enter the second phase and target tier 1 companies.

From the demand side perspective, the hypothesis is that Newton does not need a commercial partner during the first phase but the second step may involve working with partners such as consulting firms although at this moment, it is unclear if Newton will need to collaborate with 3rd parties to accelerate the adoption of the solution in more complex organizations.

The following figure contains examples of potential partnerships from the demand side perspective.
A partnership with a consulting firm is necessary to reduce the sales process and accelerate the leads-to-sale cycle by leveraging the resources and the network this types of organizations have. These sales channel partners have expertise in selling to complex organizations and can help to accelerate the adoption of the churn management solution for Tier 1 companies.

Moreover, the reputation of a leading consulting firm is fundamental to increase Newton's bargaining power in the negotiation with the TMT provider on how to split the cost savings generated by the solution and to avoid misunderstandings on how to assign the accountability of the success. Additionally, consulting firms can help Newton with the following activities:

- **Targeting companies:**
  - Focusing on TMT operators with centralized decision process, churn management part of the CEO agenda and analytics initiatives linked to top management priorities

- **Prioritizing projects:**
  - Selecting projects where it is easy to see the impact because managers expect to obtain “quick savings”
  - Delivering homogeneous projects focusing on churn management and TMT industry to generate solutions (80% standardized + 20% customized)

- **Ensuring incentives alignment:**
  - Aligning the marketing and IT managers' incentives by offering low licensing fees and a revenue sharing success model
  - Finding sponsors inside the TMT operators and partnering with marketing managers in order to deliver solutions for practical business problems
Furthermore, consulting agencies have developed a world class ability to present the benefits of a solution of this kind, in a way that marketing managers can easily understand and this is very relevant to increase the adoption rate in tier 1 companies.

The perfect timing to approach a potential sales partner is once the product is tested and validated with smaller companies and Newton has developed the strategic resources mentioned in the previous section.

Newton will “own the customer” rather than the consulting firms because the solution is measurable and generates an impact regardless of the sales channel partner. TMT players will always prefer Newton’s solution to the relationship with the consulting firms and this preference will be increased over time.

4. Final conclusion

The TMT sector is facing structural challenges due to the saturation of the telecommunications market and the increase in the competition of media and technology companies. Customers are willing to switch providers if they perceive the service does not meet with their expectations and as a consequence, TMT companies are urged to maintain their high value customers, however, few firms excel at this practice in Latin America.

In the long run, the sustainability of the TMT companies will depend on their ability to adapt to higher customer expectations because more than ever, consumers will demand greater customization and personalization from services companies.

Regarding the short run, the capability to institutionalize an experimentation culture based on a data-driven approach is a great differentiator for TMT companies and this skill is not easy for the competition to replicate.

With this context in mind, this paper presents a solution by leveraging data analytics and experimentation to develop sustainable retention discount offerings to customers in the TMT sector. This solution helps marketing managers to define business rules to allocate retention discounts for each customer based on the results of data driven algorithms and to continuously improve the rules with randomized experimentation and AB testing for a recommended retention’s discount strategy.

The main benefit of this solution is that service providers will be able to calculate the demand elasticities to different retention discounts by determining the customer’s willingness to pay for a certain service on a real-time basis. This novel capability presents huge economic and financial benefits for TMT players and therefore acquiring this technology is very relevant if not mandatory for the marketing managers.

Finally, this paper presents an urgent call to TMT executives encouraging them to adopt data driven technologies in order to create sustainable business models that enhance the customer’s experience.
5. Appendix

5.1 References and bibliography
Books:
Galit Shmueli, Peter Bruce, Mia Stephens, Nitin Patel: Data Mining for business analytics. Wiley, (2017)

Papers:
McCarthy, Clare (2009) Ovum Churn Management: The operation challenge for service providers
McKinsey Global Institute (December 2016). The age of analytics: competing in a data-driven world
Pwc (2015) Data and Analytics: Creating or destroying shareholder value? The data and analytics operating model drives the difference between leaders and laggards.

5.2 Citations

Hawkins, T. (n.d.). Retain Customers and Win “Get out there and SELL!!!