Synopsis

Customer Analytics

Domain : Customer Retention



Data Science Project

**Telecoms Focus: How to Reduce Churn and Improve Customer Retention**

By Akanksha Singh

ABSTRACT :-

Customer churn is one of the principal issues in the Telecommunications Industry. Clients massively change their specialist co-ops within the limited ability to focus time. Client Churn implies lost entire or part of the administrations from the client by any association

Every telecoms provider knows that it is much cheaper to retain an existing customer than it is to acquire a new one. However, in a highly competitive market, where one tariff change or negative customer service interaction could lose you a customer, keeping customers engaged, satisfied and, most importantly, spending on your networks is no easy task. In light of this, more and more operators are introducing personalised loyalty schemes focused on retaining their customers for longer. The truth is, loyalty and retention are probably some of the biggest challenges facing operators today, especially in the competitive global marketplace, where the current economic climate is such that customers are less willing to spend. The business of customer loyalty can be a complex decision making process. In order to achieve maximum profitability, operators must carefully mine their data and ensure they send out the most relevant targeted communications to their customer base, whilst driving efficiency and reducing costs across their business. They must also ensure they keep up with modern trends, in order to “catch their customers young” and develop lifelong relationships.

To keep these things in mind we have to spot which customer is loyal and which one going to churn so that we can make different stratergies for different groups.To identify the type of customer , we will use data science techniques.

Table of Content :

|  |  |
| --- | --- |
| Serial no. | Content |
| 1 | Business Problem Understanding |
| 2 | Business Objective |
| 3 | Advantages for Customer and company |
| 4 | Approach |
| 5 | Conclusion |

Business problem statement (GOALS) :-

1. Business Problem Understanding:

Telecom providers are finding that strong customer retention strategies because most important obstacle of telecom services is customer retention and customer churn rate.

As such, the higher your customer churn rate, the lower your chances of growing your business. Even if you have some of the best marketing campaigns in your industry, your bottom line suffers if you are losing customers at a high rate, as the cost of acquiring new customers is so high. Much has been written on the subject of the cost of [retaining customers versus acquiring customers](https://blog.kissmetrics.com/retaining-customers/), especially because study after study shows that customer acquisition costs far exceed customer retention costs. Generally, companies spend seven times more on customer acquisition than customer retention, and the average global value of a lost customer is $243. Obviously, customer churn is costly for businesses. . Client Churn implies lost entire or part of the administrations from the client by any association.

There are many factor which led to customer churn .some are :

1. Accidental

Churn due to death or shift to another country etc

1. Intentional

. Intentional churn happens when the clients change to another organization that gives comparable services, like better ideas from rivalry, further developed services and better cost for a similar service.

Retention campaigns are focused on indications that the customer is

about to leave and how to make her stay.

1. Business Objective :

companies struggle with customer retention or seem to underestimate its revenue impact.To Build Strategy which keeps engage Existing user and stops them to churn is the Objective of this project through Machine Learning model . It’s much cheaper to keep an existing customer than it is to earn a new one.Here ,we will build a model which pick out loyal customer and customers who are going to churn in near future .

By recognize both customers, by grouping them and we can build different strategy which can match their requirements so it helps to retain customer and decrease churn rate and also reduce the expending amount.When we focus on how to improve retention and reduce churn rate we will maximize the profitablility of company. If you are unable to retain your customers after you’ve just made a huge cash outlay to acquire them, they will only be a net-negative for you, unable to pay back the customer acquisition cost (CAC).

Essentially, you are only spending money to lose more money. This is why **customer retention matters to growth and profitability** and why it’s critical to focus on strategies to improve customer retention.Customer retention analytics help you determine where and [why your customers are churning](https://chattermill.com/blog/churn-analysis), and what you can do to fix this point in your [customer journey](https://chattermill.com/blog/customer-journey-map/).

Customer retention analytics are the metrics you gather and measure in order to determine how satisfied your existing customers are and what you can improve on to keep them satisfied. They measure the relationship between a customer and a brand to help predict when and why that customer might leave.

In this Project, Customer Retention analytics is used to get insights from data. Today, it’s a fact of business life, affecting processes at every level of complexity and allowing more and more companies to make data-driven decisions. In fact, [**statistics**](https://www.linkedin.com/pulse/20130604134550-284615-15-statistics-that-should-change-the-business-world-but-haven-t)show that 96% of unhappy customer don’t complain and 91% of those will simply leave and never come back.  As a result, depending on physical attributed to predict churn can be quite challenging. However, with advance in technology, business are increasingly adopting machine learning to create models that can be able to predict customer churn based on customer data.

 Businesses can leverage machine learning in reducing churn by predicting customers who are likely to leave.The classification problem uses Telco sample data to create a model that can predict customer churn based on given features. Several classification models are compared and the best is chosen for classification purposes. The target variable ‘Churn’ is a binary variable with two possible outcomes; ‘Churn’, ‘No Churn’.

1. Find the key area, gaps identified in the topic survey where the project can add value to the customers and business?

The biggest risk to telecom services is the rise in customer churn. Customer retention systems create with the help of data science will helps telecom services to make decision easy, fast and accurate in identify the of loyal customer and customer who are going to churn. Also maintain churning rate while giving right prediction about churner which helps to build right strategy to retain customers. face financial issues and not need help from other bank .The system also help to increase the revenue for company.

Particular company has been using traditional retroactive churn management based on simple statistics model.

Difficult to analysis big data with traditional method

**Churn could happen due to many different reasons and churn analysis helps to identify the cause (and timing) of this churn opening up opportunities to implement effective retention strategies.**

Customer retention analytics can be done through machine learning, allowing companies to base their product and marketing strategies on predictive customer analytics rather than less reliable predictions made manually.

Machine learning for customer retention analytics uses past customer data to predict future customer behavior. This is done using [big data](https://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/#6e7fab1913ae). In today’s data-driven world, companies can track hundreds of data points about thousands of customers. Therefore, the input data for the customer retention model could be any combination of the following:

* Customer demographics
* Membership/loyalty rewards
* Transaction/purchase history
* Email/phone call history
* Any other relevant customer data

1. Approach :

Language : Python

Algorithm : Random Forest , AdaBoosting, Decision Tree, Logistic Regression, KNeighbors ,Stacking, XGB, GradientBoosting.

Following steps are follow to make ML model.

* 1. **Dataset :**

Firstly, load dataset from file and check whether the file is loaded or not.Understand its features and check various attribute in dataset like shape of data, datatypes of dataset.

Check the columns if its datatype is correctly classified or not, if datatype is misclassified and change datatypes using astype() function. Datatypes can be classified using descriptive statistics.

* 1. **Missing Values :**

Missing values can be shown using df.isnull().sum() , heatmap etc.

Calculate the percentage of missing values , if missing values are less than 5-10% then drop records.

If data is missing for more than [50% of the observations](https://towardsdatascience.com/how-to-handle-missing-data-8646b18db0d4), it may be wise to discard it if the variable is insignificant.

Mean, mode and median is one of the most common methods of imputing values when dealing with missing data. In cases where there are a small number of missing observations, data scientists can [calculate the mean or median of the existing observations](https://towardsdatascience.com/how-to-handle-missing-data-8646b18db0d4). However, when there are many missing variables, mean or median results can result [in a loss of variation in the data.](https://blogs.oracle.com/datascience/3-methods-to-handle-missing-data) This method does not use time-series characteristics or depend on the relationship between the variables.

If data is skew we use median value to replace the null values other we use mean.

* 1. **OutLiers :**

We can we outliers using descriptive statistics or boxplot.

There are two method to treat outliers :

1. IQR Method
2. Z-score Method

If outliers are more than don’t remove it , use scaling techniques to reduce the range outliers and made data nearly normal while making model.

There are two scaling Techniques :

1. MinMax Scaling
2. Standard Scaler
   1. **Visualization :**

Visualization is useful to gain meaning-ful insights from data which is haed to get by looking whole data.

Through visualization , we firstly analyse univariate data then bi-variate and multi-variate .

As we know ,data are of two type :

Qualitative

Quantitaive

So , we uses different types of graphs suited for different variables such as pie chart, bar graph, scatter plot, heatmap,lmplot,boxplot, etc.

* 1. **Separate variables :**

Target variable is separated from independent variables,which will in numeric form.

Independent variable are present in qualitative and quantitative mode , so we scale the numeric columns if they have large value on same scale which will be helpful in making model.

We encode qualitative data because machine learns from numbers through various encoding method.

* 1. **Train-Test :**

After creating Independent varibles dataframe(X) and Dependent variable(y) , we split the data into training and testing dataset so that model will learn from training dataset and model will be tested on testing set. If the data has imbalanced data then use resampling techniques to made data balanced.

* 1. **Algorithm :**

Choose algorithm which you want to use for building model.

Used algorithm in this project

1. **Random Forest**

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning,** which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.

**Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.**

**The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.**

**Why use Random Forest?**

It takes less training time as compared to other algorithms.

It predicts output with high accuracy, even for the large dataset it runs efficiently.

It can also maintain accuracy when a large proportion of data is missing.

1. **Adaboosting**

AdaBoost can be used to boost the performance of any machine learning algorithm. It is best used with weak learners. These are models that achieve accuracy just above random chance on a classification problem.

Weak models are added sequentially, trained using the weighted training data.The process continues until a pre-set number of weak learners have been created (a user parameter) or no further improvement can be made on the training dataset.

Steps follows by Adaptive Boosting :

1. Build a model and make predictions.
2. Assign higher weights to miss-classified points.
3. Build next model.
4. Repeat steps 3 and 4.
5. Make a final model using the weighted average of individual models.
6. **Decision Tree**

Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome. In a Decision tree, there are two nodes, which are the Decision Node and Leaf Node. Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches.

In order to build a tree, we use the CART algorithm, which stands for Classification and Regression Tree algorithm.

A decision tree simply asks a question, and based on the answer (Yes/No), it further split the tree into subtrees.

1. **K-NearestNeighbors (KNN)**

K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique.

K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.

K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm.

K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems.

K-NN is a non-parametric algorithm, which means it does not make any assumption on underlying data.

1. **Logistic Regression**

Logistic regression is a process of modeling the probability of a discrete outcome given an input variable. The most common [logistic regression models](https://www.sciencedirect.com/topics/computer-science/logistic-regression-model) a binary outcome; something that can take two values such as true/false, yes/no, and so on.

Logistic regression is a simple and more efficient method for binary and linear classification problems. It is a classification model, which is very easy to realize and achieves very good performance with linearly separable classes.

1. **Stacking :**

Stacking is a way to ensemble multiple classifications or regression model.  The point of stacking is to explore a space of different models for the same problem. The idea is that you can attack a learning problem with different types of models which are capable to learn some part of the problem, but not the whole space of the problem. So, you can build multiple different learners and you use them to build an intermediate prediction, one prediction for each learned model. Then you add a new model which learns from the intermediate predictions the same target.  
This final model is said to be stacked on the top of the others.

1. **Gradient Boosting :**

gradient boosting is one of the boosting algorithms it is used to minimize bias error of the model. Gradient boosting algorithm can be used for predicting not only continuous target variable (as a Regressor) but also categorical target variable (as a Classifier). When it is used as a regressor, the cost function is Mean Square Error (MSE) and when it is used as a classifier then the cost function is Log loss.

1. **XGBoosting :**

XGBoost stands for e**X**treme **G**radient **B**oosting.XGBoost is an algorithm that has recently been dominating applied machine learning.

XGBoost has proved its mettle in terms of performance – and speed.XGBoost or extreme gradient boosting is one of the well-known [gradient boosting](https://analyticsindiamag.com/gradient-descent-everything-you-need-to-know-with-implementation-in-python/)techniques(ensemble) .

**Features of XGBoost:**

1. Can be run on both single and distributed systems(Hadoop, Spark).
2. XGBoost is used in supervised learning(regression and classification problems).
3. Supports parallel processing.
4. Cache optimization.
5. Efficient memory management for large datasets exceeding RAM.
6. Has a variety of regularizations which helps in reducing overfitting.
7. Auto tree pruning – Decision tree will not grow further after certain limits internally.
8. Can handle missing values.
9. Has inbuilt Cross-Validation.
10. Takes care of outliers to some extent.
    1. **Parameter Selection :**

We use Gridsearchcv to search for best parameter than we again build model using best parameter.

* 1. **Evaluation :**

After building of model, use various metrics to compare model which we can deploy in production.Metrics such as accuracy score, ROC\_AUC score , False negative or False Positive from confusion matrix ,F1-score etc.

We can so

1. Conclusions :

After Building the model, we deploy in production . we can predict the future action of customer and based on this prediction we build various strategy to retain customer .This will Boosting our Profitablity and decision making process.

Model will help us to work on Significant features more than insignificant features. Focus on customers who are actually at risk and improve their experience.

Save up on acquisition cost to replace the Churning customers.

Having stable revenue helps you to grow and acquire new customers.

References : Kaggle ,Youtube.