Churn Analysis Model

Overview

Churn refers to the loss of customers or clients from a company's customer base. A churned customer is a person who has stopped doing business with a company. This can happen for a variety of reasons such as poor service, dissatisfaction with products, or simply because the customer found a better alternative. Churn is a significant challenge for companies because it can impact their revenue, customer base, and overall growth. To prevent churn, companies often focus on improving customer satisfaction, offering incentives to retain customers, and addressing any issues that may be causing customers to leave.

So the Churn Analysis Model is a comptinot of three different models .

1. Churn definition model: the goal of this model is to create a systmitic way that can be used to define the customer based on there purchase history, for example the customer that stop purchase the company products and services for 90 days can be consider a churn , the role that needed to be define for the churn customer can be vary significantly based on the company product and services nature .
2. Churn predictive model : this is a machine learning classification model that takes the customer data set and the output form Churn definition model as input to create a model that can identify the possibility for a customer to become churn.
3. Churn descriptive model: is a interactive model that shows the churn rate for a company along with different customer demographic and attributes , which will give the company more insightful ideas on how to prevent the lost cousterm or the potential churn customer form leaving the company .

The integration of the three models will create a good Churn Analysis Model that can give the clint a detailed view at the customers loyalty and satisfaction.

The model specification

1. For the Churn definition model we didnot have the data to create the model since the dataset in hand has already dealt with this stpe.

2. The phases to create the Churn predictive model. was the following:

Gather and Preprocess Data

The data used for creating the model can be found at this link, the dataset proceed by converting all the numerical columns from object to numerical , and convierting the gender clounm to 0 and 1 to be easier to interpret and work with.

Feature Engineering

To work with this dataset ,we needed to conifierm there is no dependency between the column (no clounm dircitly derived from another) this can be done using the Chi Square Test .after confirming the Independence ., the feature Engineering can be done using different method , for this model we used Correlation Matrix between the targeted column ‘Exited’ and the rest of the dataframe , we also used Univariate Feature Selection method, and we found that the most important features to consider are *'Age', 'Balance', 'EstimatedSalary' , 'IsActiveMember'* .

Model creating and evalouiton

Since this a classification problem we used the five different classification algorithm , the following table shows the performance for each one of them when we train them on 80% of the data and test them on the remaining 20%

Model	Parameters	Accuracy	Remark
Logistic regrension	Default	79.75%	Useless ,set all to therefore it got high soccer since the most of customer are not churn
Random forest	Default	81.00%	Highest score adn best performance
SVM	Default	79.75%	Useless ,set all to therefore it got high soccer since the most of customer are not churn
Naive Bayes	Default	78.10%	Efficeted by the imbalance between the representation of the two class
KNN	Default	76.40%	Efficeted by the imbalance between the representation of the two class

Model Maintenance and Improvement

There are very big opportunity to improve the model that will make it more efficient such as :

- Fune tuning the parameter for each model
- Concieder using deep learning models
- Dealing with the imbalance at the dataset.
- Training more feature selection algorithms.

Useful reference : Churn-Modelling

3. The phases to create the Churn descriptive model. was the following:

Gather and Preprocess Data

The data used for creating the model can be found at this link and original obtained form **here** . to create the model we add two more columns to the dataset, count clounms , churn flag (1 if chrun 0 if not) , also we calculate the churn rate (count cloumn / churn flag). This add cloumns and churn rate will be very benficial to create the churn descriptive model.

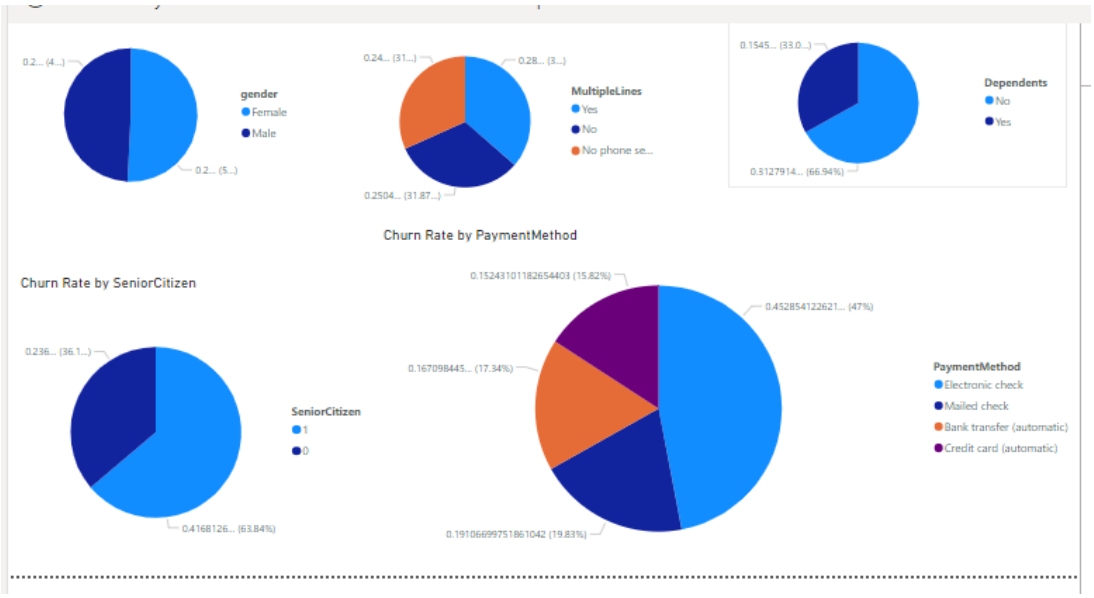
Feature Engineering

The dataset contian many clounms , almost all of them can be useful to create the churn descriptive model , however , for this model we used SeniorCitizen, gender ,payment method ,dependents,monthly charge and multil line columns , monthly charge is a continuous therefore for the smplisty we round it and we also created another categorical version of it using the following DAX formale :

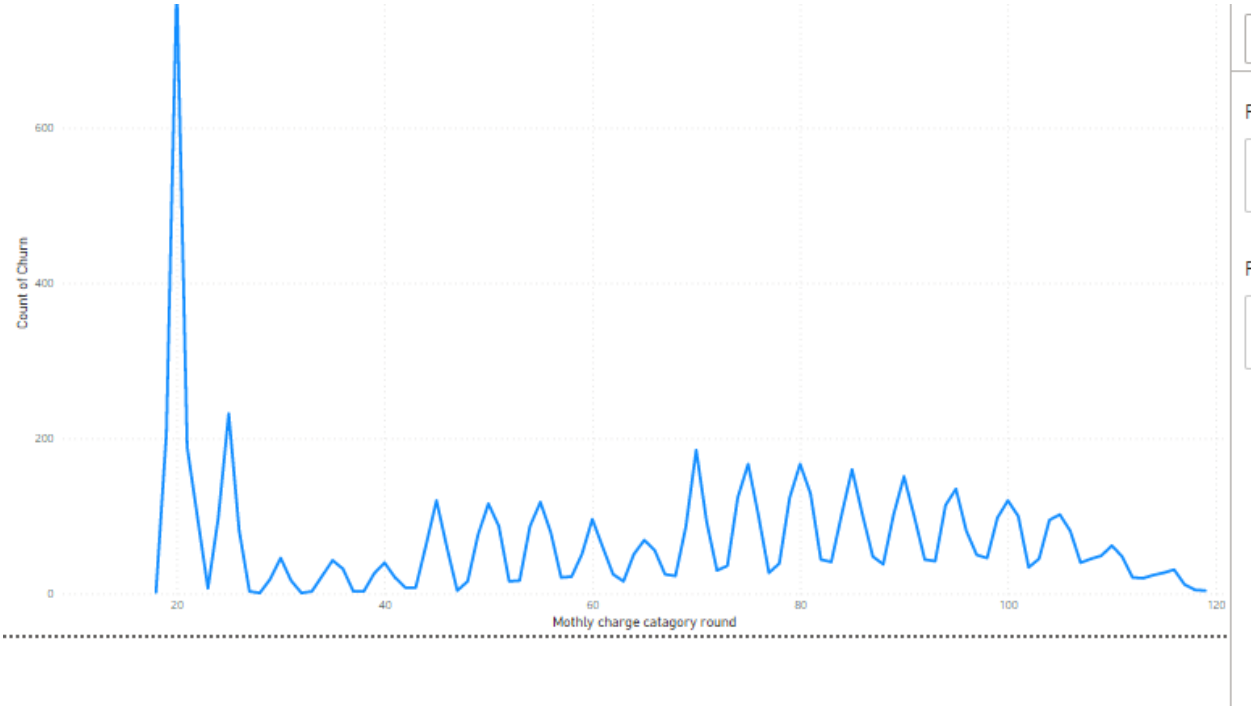
```
MCas cataigoriy = IF([MonthlyCharges]<=28,"0-
28",IF([MonthlyCharges]<=43,"28-43",IF([MonthlyCharges]<=68,"43-
68","+68")))
```

Model creating

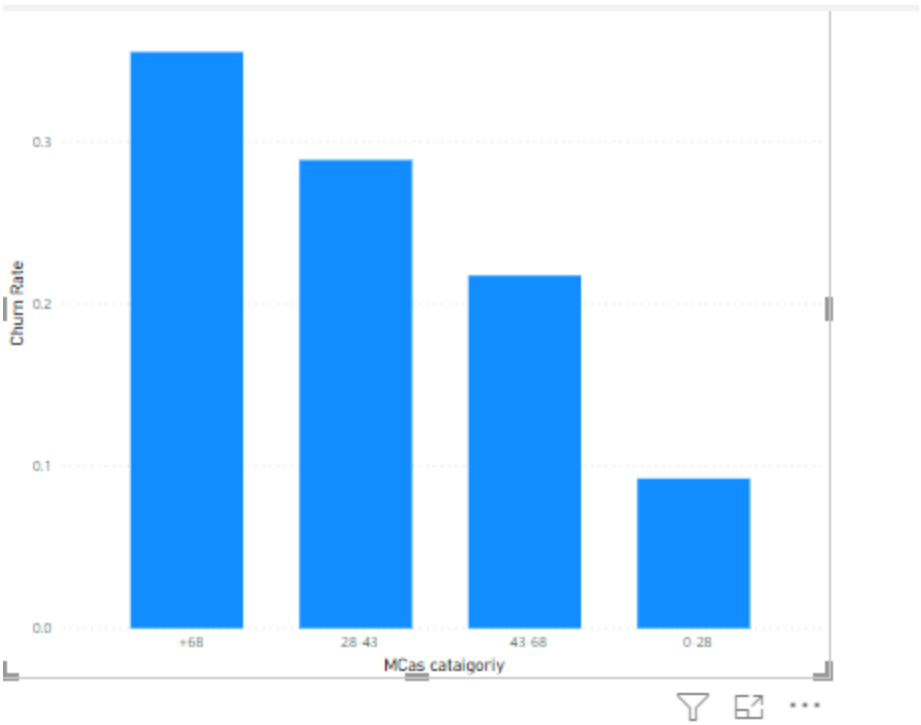
The model has three page :
Page 1 : shows the churn rate and the distrution of the churn with the other customer attributes and demographics. Since the visualizations are interactive we can show many different combinations between the different attributes.



Page 2 : shows how can we work with a column contain continuous data, the visualizations shows the the count of the churn and the rounded monthly charge.



Page 3 : shows the count of the churn and the monthly charge as categorical column .



Model Maintenance and Improvement

There are very big opportunity to improve the model that will make it more efficient and more real world like .

1. Real world : using real world data will be a very good way to exprice the needed effort for data processing and feature engineering , it may consume more time but it will be very beneficial.
2. Sql Dataset : one way to be praberd for more realistic clints needs ,is to create the same dashboard unig Sql dataset, this will involve on two important practices, extract the dataset from the Sql server using a connetion btween the dataset and python , using the python code to load the data to Power Bi
3. More insightful and complex result : the created dashboard is good to give a general idea on the Customer-segmentation concept , however , the model is quite premivite and does not give more in depth insight about the data, this may be caused by the dataset in hand , however we need to create more insightful dashboard .
4. Creating a shareable dashboard , which mean to ues power Bi serve services to create live updated and shareable dashboard.
5. Improve the current visualization : this can be done by referring to visual analytic concepts , where it shows which type of graphs to ues for which

data type, which color to use , what type of KPI and card to create and more efficient graphs properties.

6. Create a list of actionbale procedure that can be executed by the clint in order to make the most of the churn coustmers .