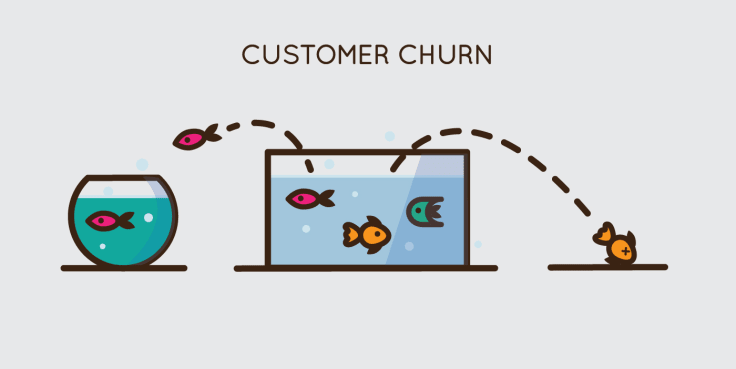
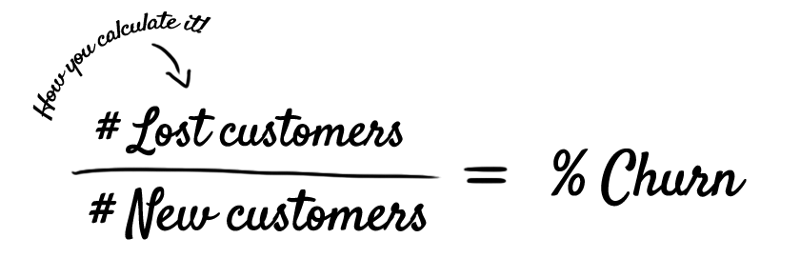
Telecom Customer Churn Prediction

Image Source:[https://www.livechat.com/success/churn-rate/](https://www.livechat.com/success/churn-rate/" \t "_blank)

**What is Customer Churn and Churn Rate ?**

Customer churn is nothing but the proportion of the customer lost in the given period of time .It can normally be a year. Churn rate is a metric that shows the percentage of customers you lose when their subscription expires or they stop using your product. Calculating customer churn rate is crucial to check the health rate of your business.

Image Source: <http://customer-churn.com/>

Customer are the important assets for any company ,Company start facing loss when these customer leaves the company ,Customer can be loss because of the competition company giving better offer than the last one or it can be many other reasons too. Hence, here comes **Customer churn rate** into picture which is one of the important metrics for companies to evaluate their performance.

Let us understand this with an example, If company had 400 customers at the beginning of the month and only 360 at the end of the month, means company’s churn rate is 10%, because company lost 10% of the customer base. Companies always try to minimize the churn rate to as close as 0%.

Here, IBM provided customer data for Telecom industry to predict churn customer based on demographic, usage and account based information. Main objective here is to analyze churn customers’ behavior and develop strategies to increase customer retention.

Source :[https://raw.githubusercontent.com/dsrscientist/DSData/master/Telecom\_customer\_churn.csv](https://raw.githubusercontent.com/dsrscientist/DSData/master/Telecom_customer_churn.csv" \t "_blank)

Dataset has information related to,

**Population Tally:**

**Gender**— Male / Female

**Age range** — In terms of Partner, Dependent and Senior Citizen

**Services:**

**Phone service**— If customer has Phone service, then services related to Phone like Multi-line Phone service

**Internet Servic**e — If customer has Internet service, then services related to Internet like Online security, Online backup, Device protection, Tech support, Streaming TV, Streaming Movies

**Account type:**

**Tenure** — >How long customer is with the company?

**Contract type** — >What kind of contract they have with a company? Like Monthly bases, On going bases — If on going bases, then One month contract or Two year contract

**Paperless billing** —> Customer is paperless billion option or not?

**Payment method** — >What kind of payment method customer has? Mailed check, Electronic check, Credit card (Automatic), Bank transfer (Automatic)

**Usage:**

 Monthly charges

 Total charges

**Target:**

**Churn**— >Whether customer left the company or still with the company?

**Problem Description**

**Why customers are leaving the company?**

The reasons behind the customer leaving company could be

*→ High charges*

*→ Better offer from competitor*

*→ Poor customer service*

*→ Some unknown reasons*

**How to detect the churn customers?**

*→ Monitoring usage*

*→ Analyzing complains*

*→ Analyzing competitors offers*

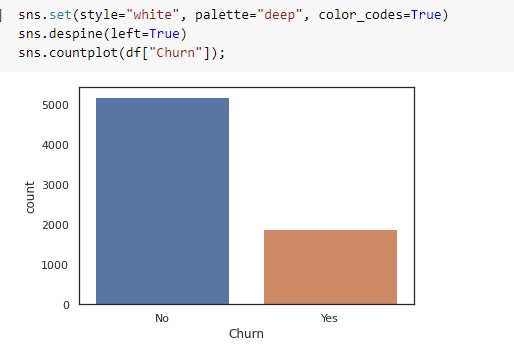
**How to prevent customers from leaving a company?**

Once you detect high risk customers, apply

*→ Retention plans*

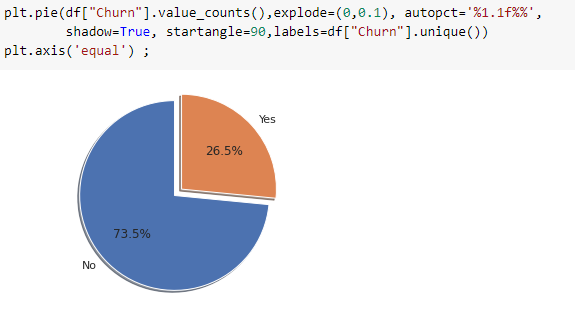
*→ Improve customer service*

**Exploratory Data Analysis**



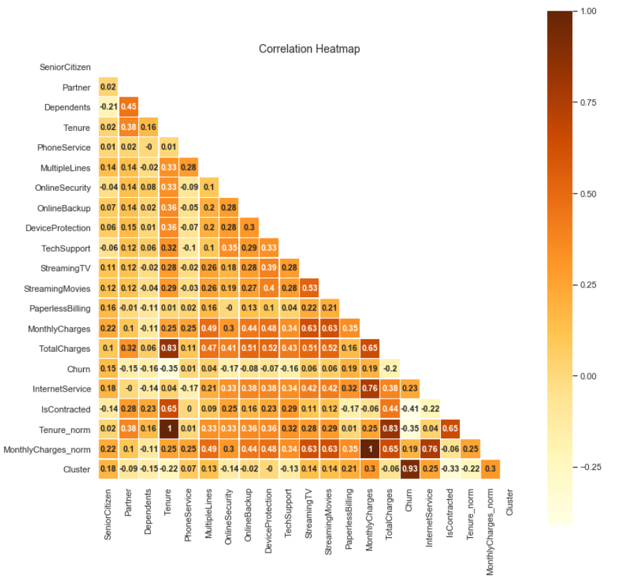
Using the code above, we can clearly see the churn rate ,That in Bar graph ,we can see that how many customer has left the business and the count is near 2000.

Let us see another chart:



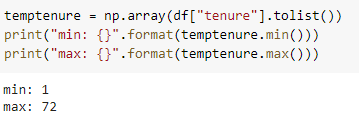
Analysis shows that ***Churn rate of the Telecom company is around 26.5%.***

**Correlation with all the features**

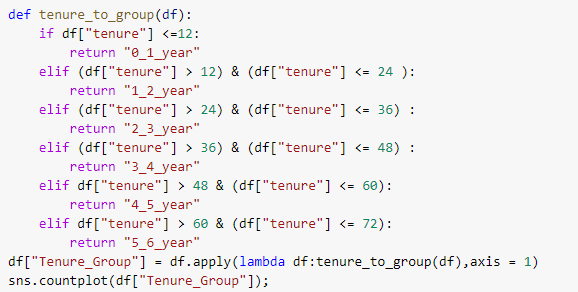


From correlation matrix we can clearly see, That the features like Tenure, Monthly charges and Total charges are highly correlated with services like Multiple Phone Lines services and Internet services like Online Security, Online Backup, Device Protection, Tech Support, Streaming TV and Streaming Movies services.

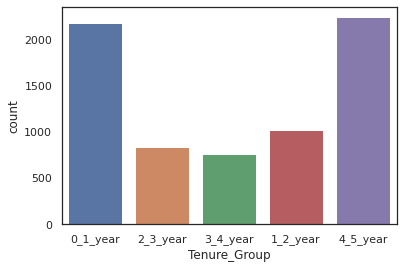
**Distribution of Churn Rate in Tenure**



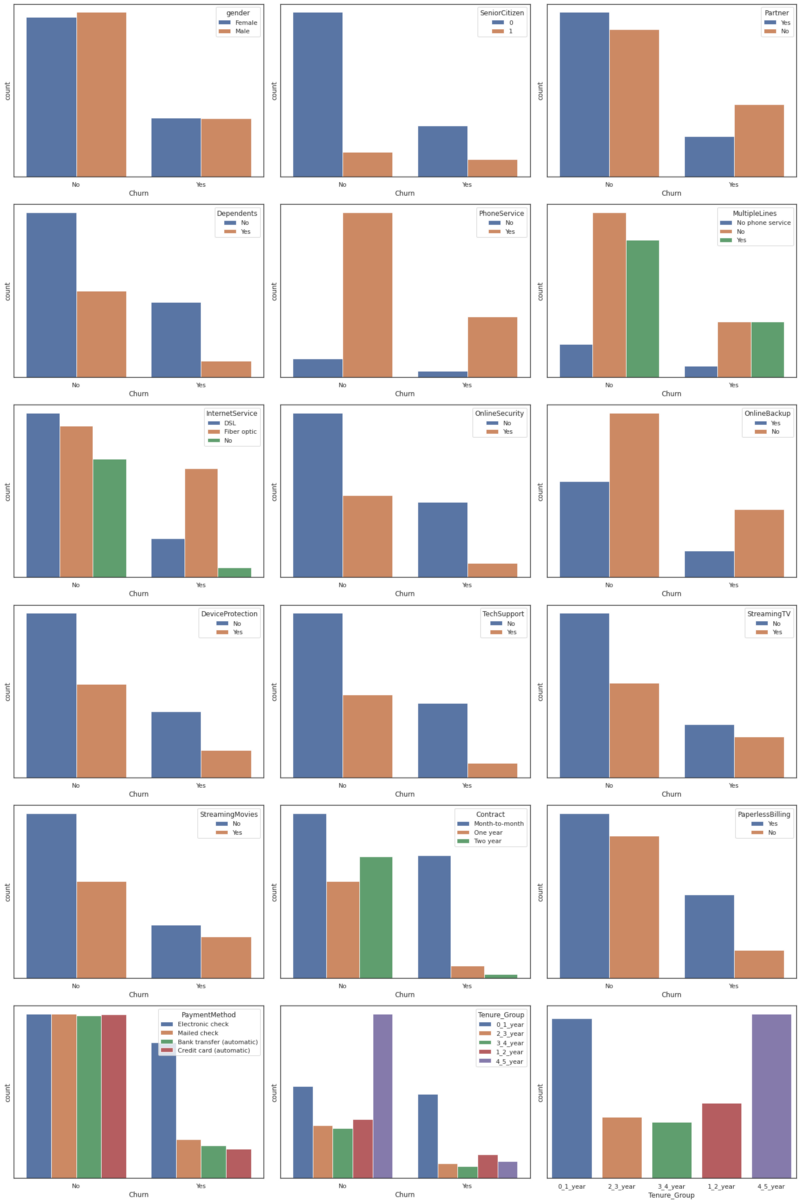
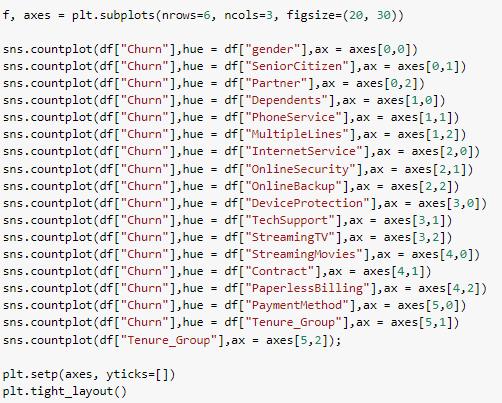
from above code , we can understand the period of time the business have evolved .



To check the distribution of churn rate over the period of time we have grouped all the tenure together in one graph



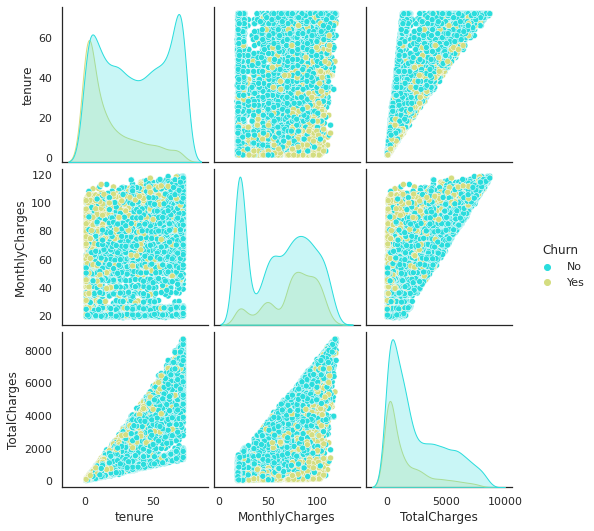
**Distribution of Churn Rate in all the Features in the given Dataset**



From above analysis we can clearly see that churn customers are likely to

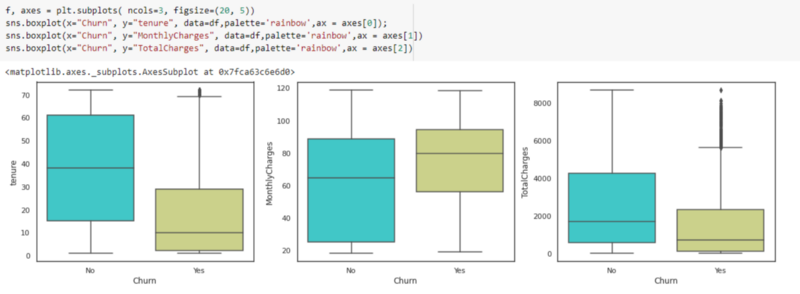
* *not have Partner and Dependents; Meaning likely to be single.*
* *have Internet service and specifically Fiber optics*
* *not have online security service, online backup service, device protection service, Tech support service*
* *have streaming TV and streaming Movies services*
* *be with monthly based plan*
* *have paperless billing service*
* *have electronic check payment method***s**

**Distribution of Churn in Tenure, Monthly Charges and Total Charges**

https://cdn-images-1.medium.com/max/800/1*FjBLNGN_sh95SxAytwZrUg.png

We can clearly see the customers are more likely to leave over the tenure and there is drastic change over the period of time because of monthly charges rate

Let us confirm that the results we are getting is correct or not by using different method



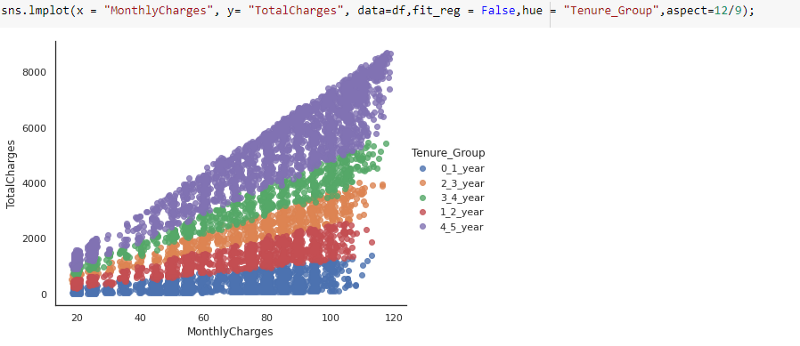
By the above plotting, we can clearly see that the customer started leaving the company over the period of time

From distribution, churn subscribers are

*→* more likely to leave company who’s tenure is less than a year

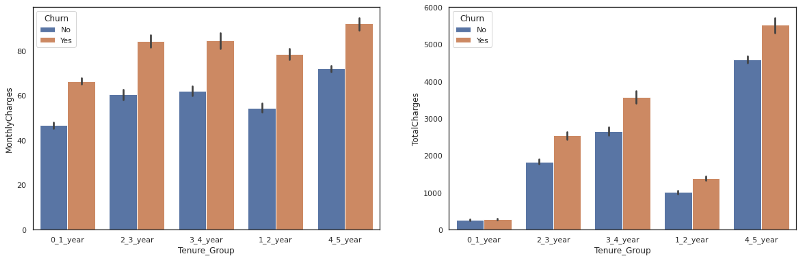
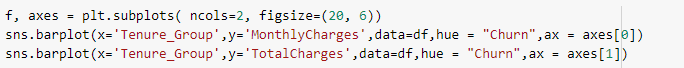
*→* more likely to have more than $65 monthly charges

**Monthly Charges Vs Total charges**



Result shows that Churn customers have ***more Monthly charges*** compared to existing customers. Also, there is a drastic difference in **Tenure** and **Total Charges** for Churn vs existing customers.

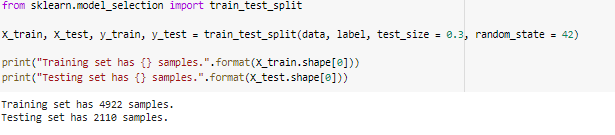
**Monthly Charges Vs Tenure and Total Charges Vs Tenure**



Here, we can see if the **tenure is low, low monthly charges** for customer like in Month-to-month contract plan and when the **tenure is low , High Monthly charges** for customer like in Paperless billing, Month-to-month contract plan and when the **tenure is High, High Monthly charges** for customer like in Paperless billing, One/Two year contract type also, we can see there is a drastic change in **Total charges and Tenure**

**Data Preprocessing**

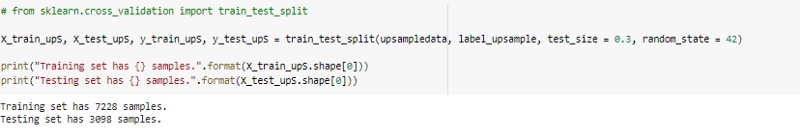
Splitting the datasets into two group → Train and Test



There is a imbalance between the original data and the training

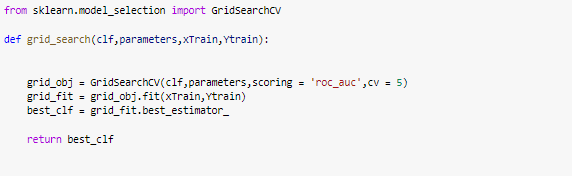
Original data sets has 7228 samples but here in the training sample it takes only 4922 .So, to rectify this imbalance we are going to use Unsampling Method.

Let us try the Unsampling Method:



**Hyper-parameter tuning**

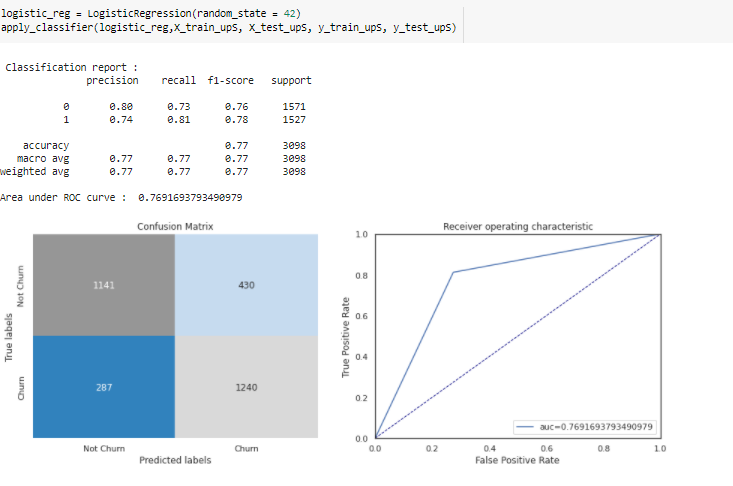
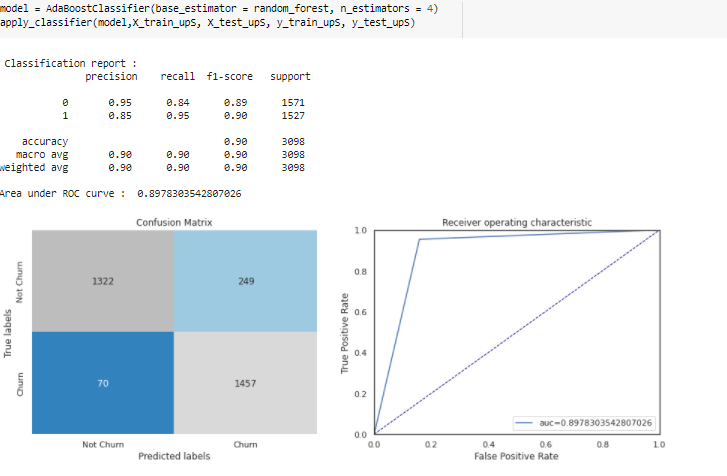
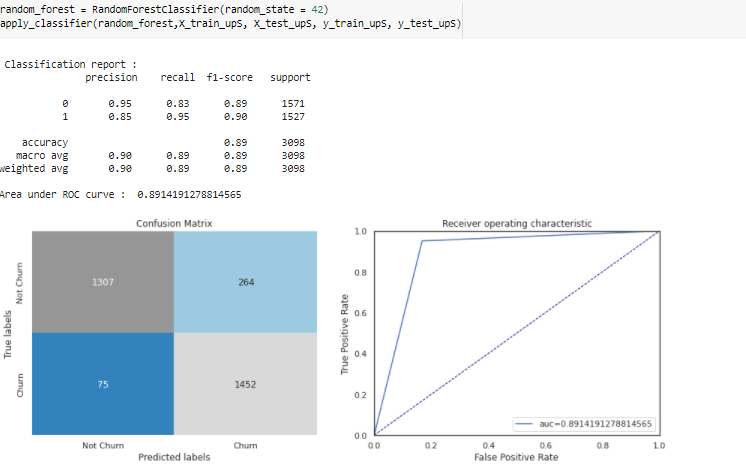
Using GridSearchCv() method, Let us identify the best parameters for our models



**Model Selection**

 → Comparing models like Random forest Classifier, Logistic regression, Decision Tree Classifier , Ada Boost Classifier and many more with AUC curve

 →Measuring scores like Accuracy, Precision, Recall and F1 metrics



Using scores like Accuracy, Precision, Recall and F1 metrics finding the best model

**Model conclusion**

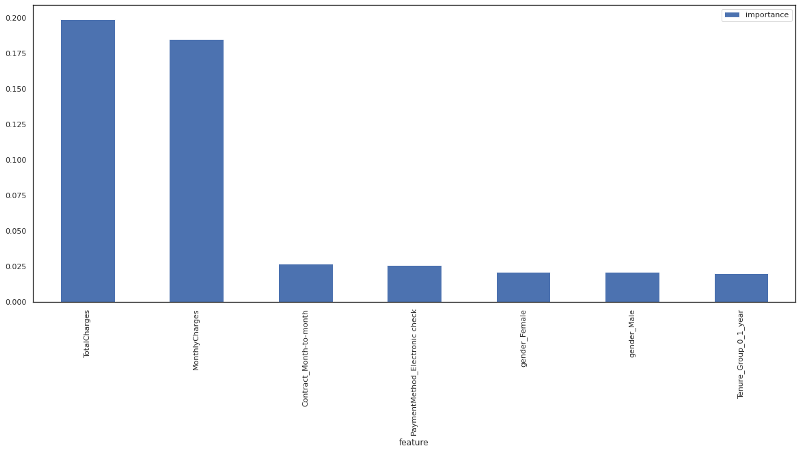
Using the model comparison process and using the scores of F1 we have identify the best model ,up sampling data works better during training process.

From above analysis, Ada Boost Classifier with original dataset has stable and best score. So, I have used it for prediction process.

Ada Boost Classifier model suggested important features like

* Total charges, Tenure, Monthly charges, Contract type, Payment method, Internet service type, Paperless billing

Most of them, we already analyzed during our EDA process.



**Conclusion**

In this project, I have tried to divide the customer churn problem into method like Exploration, Data cleaning and Visualization, Churn Rate prediction, Data Preprocessing, Model Selection and Saving of Best model.

Based on this analysis, we can help Data Maintenance team to analyze high risk churn customers before they leave the company.

However, we can even use different data sources like customer inquires, seasonality in sales, more Population Tally information to make our prediction more accurate.

If you want to get more details above the code, I am attaching the link below.

<https://github.com/shailzafiroz/Evaluation_Project/blob/main/ChurnDT.ipynb>