A

Capstone Project Final
Report On

"Customer Retention Retail"

Presented By
Mr. Sam Darshan M,
Mr. Raghul kumaran,
Mr. Baipalli Prem kumar Eswar,
Mr. Christopher X,
Mr. Anjan Chakraborty

Under the Guidance of
Mr. Jatin Bedi

GREAT LAKES INSTITUTE OF MANAGEMENT AND GREAT LEARNING

BATCH: PGPDSE-FT ONLINE MAR22

INDEX

Industrial Review:

Many review studies were handled to provide valuable insights into customer retention issues and factors that could influence it positively and effectively.

Generally, it is recognized that there is a positive relationship between customer retention and profitability. Customer retention enables the company to increase profitability and revenue. Thus, the small increase in customer retention could have a positive impact on profitability. Indicated that customer retention indicates customer's intention to repurchase a service or a product from the service provider. The customer retention is defined as the future propensity of a customer to stay with the service provider. It should be a continuous process to find a customer and retain them in a long-term relationship.

Customer retention survives when the companies can fulfil customer expectations and additionally maintain it in long-term relationships to ensure long-term buying decisions.

The topic of customer retention is argued in business economics commonly within the perspective of relationship marketing, which considers customer relationships as one of the primary concerns with the long-term objective of developing and maintaining them.

Companies in a variety of sectors have increasingly started managing customer churn proactively, generally by detecting customers at the highest risk of churning and targeting retention efforts towards them (Ascarza, 2018) [1]. While there is a vast literature on developing churn prediction models that identify customers at the highest risk of churning, no research has investigated whether it is indeed optimal to target those individuals. Combining two field experiments with machine learning techniques, the author demonstrates that customers identified as having the highest risk of churning are not necessarily the best targets for proactive churn programs. This finding is not only contrary to common wisdom but also suggests that retention programs are sometimes futile not because firms offer the wrong incentives but because they do not apply the right targeting rules. Accordingly, firms should focus their modelling efforts on identifying the observed heterogeneity in response to the intervention and to target customers on the basis of their sensitivity to the intervention, regardless of their risk of churning. This approach is empirically demonstrated to be significantly more effective than the standard practice of targeting customers with the highest risk of churning

Customer experience plays a key role for firms in creating a sustainable competitive advantage and building good customer relationships (Andreini et al., 2018) and is thus a fundamental element of firms' success in both online and offline channels (Barari et al., 2020; Bleier et al., 2019; Bustamante & Rubio, 2017).

Literature Review

Customer retention survives when the companies can fulfil customer expectations and additionally maintain it in long-term relationships to ensure long-term buying decisions [6-8]. The topic of customer retention is argued in business economics commonly within the perspective of relationship marketing, which considers customer relationships as one of the primary concerns with the long-term objective of developing and maintaining them [9-11].

Many previous studies indicated that companies should always manage customer satisfaction to achieve the retention stage. According to [12] "satisfaction is an overall customer attitude towards a service provider". In [13], authors added that satisfaction is an emotional reaction regarding what customers expect and what they receive, including the fulfilment of needs and goals. Customer retention states a desired outcome in the future to satisfaction, so long-term of relationship is demonstrated by satisfaction. Although customer satisfaction does not guarantee repurchase, it still plays a vital role in ensuring customer retention. While many studies on customer retention had long focused on customer satisfaction, additional factors are stated as an influence in customer retention, such as trust and commitment, in "The Commitment-Trust Theory of Relationship Marketing," which is the most influential Relationship Marketing, suggests that the centre of successful relationship marketing is the relationship of commitment and trust. They urged the importance of commitment and trust that leads to build a positive correlation between company and customers and encourage efficiency, productivity, and effectiveness. The degree of trust between service provider and customer is significantly influenced by the quality of the service, which results in an effective commitment to the provider, and Factors Affect Customer Retention: A Systematic Review 657 enhancing commitment is important since it leads to an intention to invest further and reinforce the relationship with the provider.

Data set and Domain

Data Dictionary

Column Name	Data Type	Definition
custid	object	Computer generated ID to identify customers throughout the database
retained	int64	1, if customer is assumed to be Retained , 0 = Not Retained
created	datetime64	Date when the contact was created in the database - when the customer joined
firstorder	object	Date when the customer placed first order
lastorder	object	Date when the customer placed last order
esent	int64	Number of emails sent
eopenrate	float64	Number of emails opened divided by number of emails sent
eclickrate	float64	Number of emails clicked divided by number of emails sent
avgorder	float64	Average order size for the customer
ordfreq	float64	Number of orders divided by customer tenure
paperless	int64	1 if customer subscribed for paperless communication (only online)
refill	int64	1 if customer subscribed for automatic refill
doorstep	int64	1 if customer subscribed for doorstep delivery
favday	object	Customer's favourite delivery day
city	object	City where the customer resides in

Variable categorization (count of numeric and categorical)

Number of Numerical columns - 11

Numerical column names - retained, esent, eopenrate, eclickrate, avgorder, ordfreq , paperless, refill, doorstep, create_first, first_last

Number of Categorical columns - 2

Categorical column names - favday, city

Pre-Processing Data Analysis (count of missing/ null values, redundant columns, etc.)

The data has been taken from kaggle.com**.** The dataset indicates the overall purchase history of the customers who have purchased from the Online Tea Retail Store. The Data set consists of different fields which might consist of different parameters which affect the customer churn rate.

The online tea retail store data considered here is from 2008 and 2018 which consist of around 30801 observations and has 15 columns. The online tea retail store sells tea of different flavours across various cities in India. The dataset contains data about the store's customers, their orders, quantity ordered, order frequency, city, etc..

The Dataset consists of some missing values. The Columns which has around 20 missing values include Custid, Created, firstorder and lastorder. The column Custid contains the serial number of the customer, which is redundant for further analysis. Thus, we drop the column. We have Dropped a few observations where the format was not appropriate (''1/0/00', '00:00:00'). Since the Null value count was very negligible as opposed the count of total observation we have dropped those as well.

The columns firstorder and lastorder had datatypes(Object) that were inappropriate and hence we have changed the datatype(Datetime64). We also observed that the data is skewed and not normal. Hence, we will be using Transformation techniques to treat that.

We also observed an imbalance in the data where the retained customers consisted of 79.46% and churned customers consisted of 20.54%. We have observed that few of the columns had outliers and we have used the IQR technique to impute the outliers using Boxplot.

Alternate sources of data that can supplement the core dataset (at least 2-3 columns) -

We did not find any such data that can supplement the core dataset.

Project Justification -

Project Statement -

This Online Tea Retail store has encountered the problem of losing out existing customers in recent times. The Business owner isn’t aware of the reasons why the customers are churning and would like to bring down the churn rate.The churn rate, also known as the rate of attrition or customer churn, is the rate at which customers stop doing business with an entity.

Modern businesses nowadays employ complex algorithms to predict customers that are most likely to churn, i.e. move away from the company. By using such algorithms, companies can know in advance the customers that are most likely to give up the company’s services and therefore, come up with customer retention strategies to mitigate the losses that the company might face. In this project, you will analyze customer-level data of Online Tea Retail, build predictive models to identify customers at high risk of churn, and identify the main indicators of churn.

Some of the customer retention strategies can include sending weekly email newsletters, collecting customer feedback, Building a community of Tea connoisseurs on social media, and Rewarding the promoters and loyal customers. By implementing the above-stated approaches, the store can increase its Customer retention rate and thereby increase sales without incurring a huge customer acquisition cost.

Currently, business owners go with their gut feel and do not have a clear idea about the existing customers. As a result, the business owners end up targeting the wrong customer base to improve retention. Hence the business owners lose a substantial amount of time and money in the process.

We can arrive at a conclusion that customer retention makes the business more sustainable, and reduces customer acquisition costs. Customer retention saves you cash and cuts back your selling expenses by keeping your recent customers who are already familiar with your products and services. Thereby it helps the business to make more profits.

Complexity involved -

There can be a lot of factors that influence the customer behaviour. Since these factors are all dynamic the conclusions that we may arrive at may not be relevant in the future. Continuous training and monitoring of ML models in production is important compared to static validation and testing techniques in order to overcome this issue.

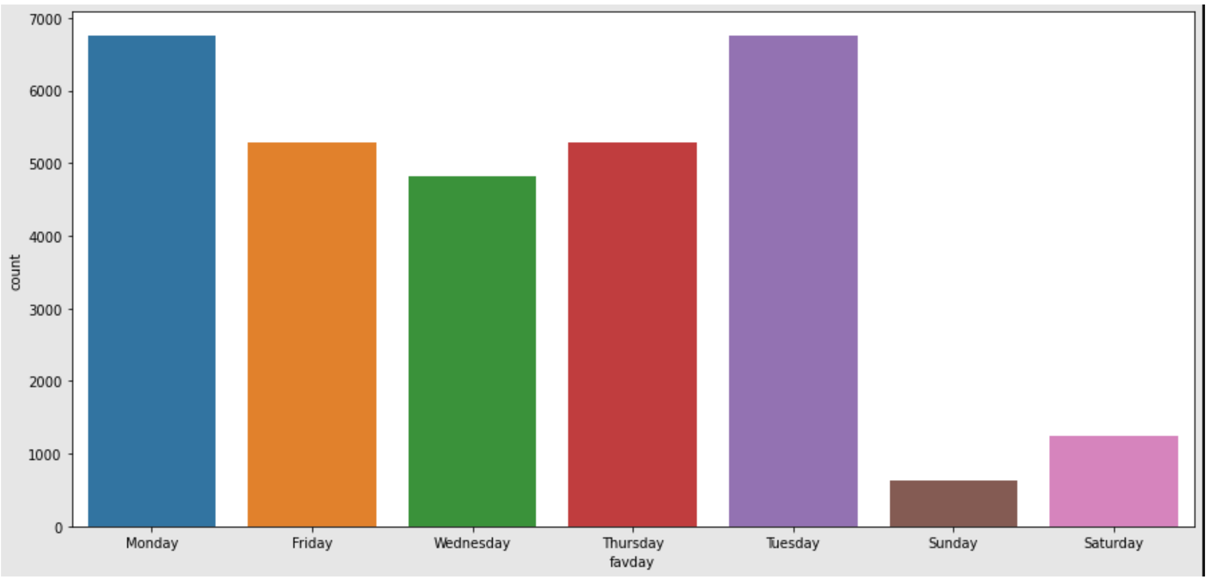
It is important to build a new system from scratch with updated data. There is also a need to consider a retraining strategy and come up with a clear understanding of the business objectives relevant to that particular time.

Project Outcome Commercial -

- To Increase customer retention by adopting different strategies
- To come up with an effective targeted marketing strategy to avoid customer from churning
- To formulate strategies for demographics of different locations.
- To understand customer behaviour in deep manner.

EXPLORATORY DATA ANALYSIS (EDA)

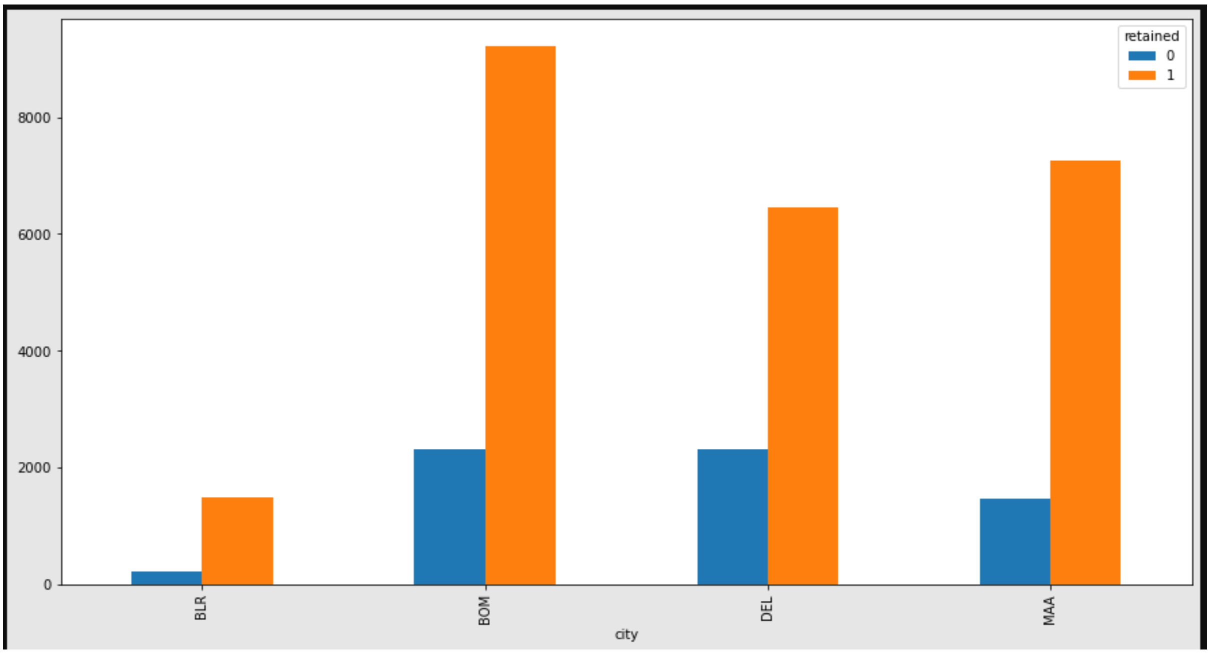
We have plotted **a bar plot on favday** to understand the trends of delivery and orders across the different days of the week.



Inferences-

- The number of orders delivered are highest on Monday and Tuesday since the number of orders placed are maximum on Saturday and Sunday.

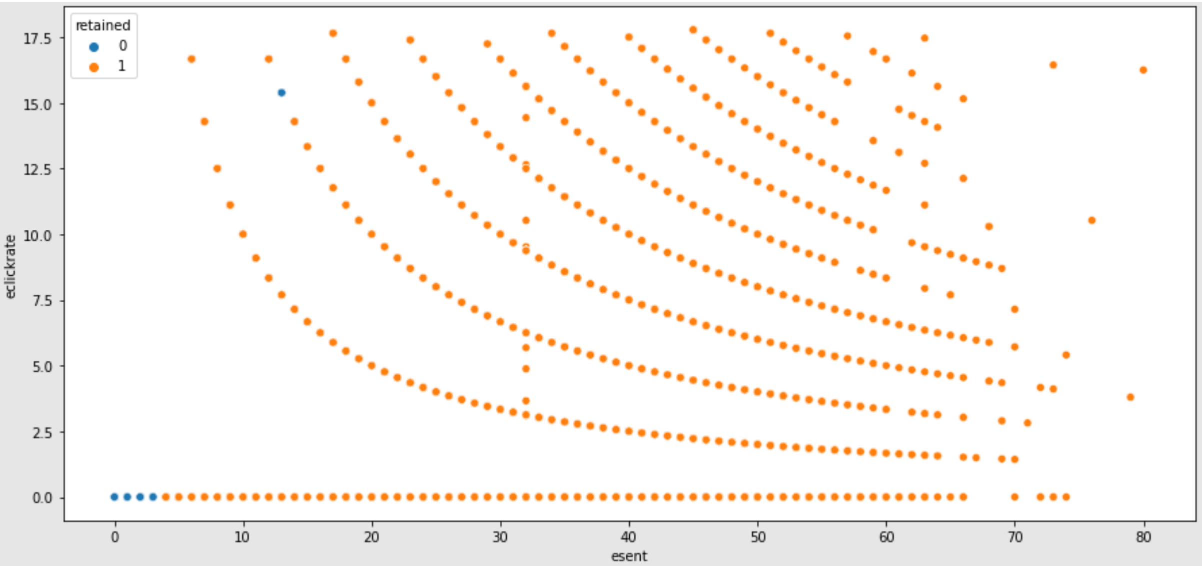
We have plotted a **bar plot on orders across different cities** to find the retention rate in different cities



Inferences -

- The retention rate is maximum in Bombay because of high population and rainy climatic conditions.
- Bangalore has the least retention rate because people in Bangalore prefer coffee over tea.

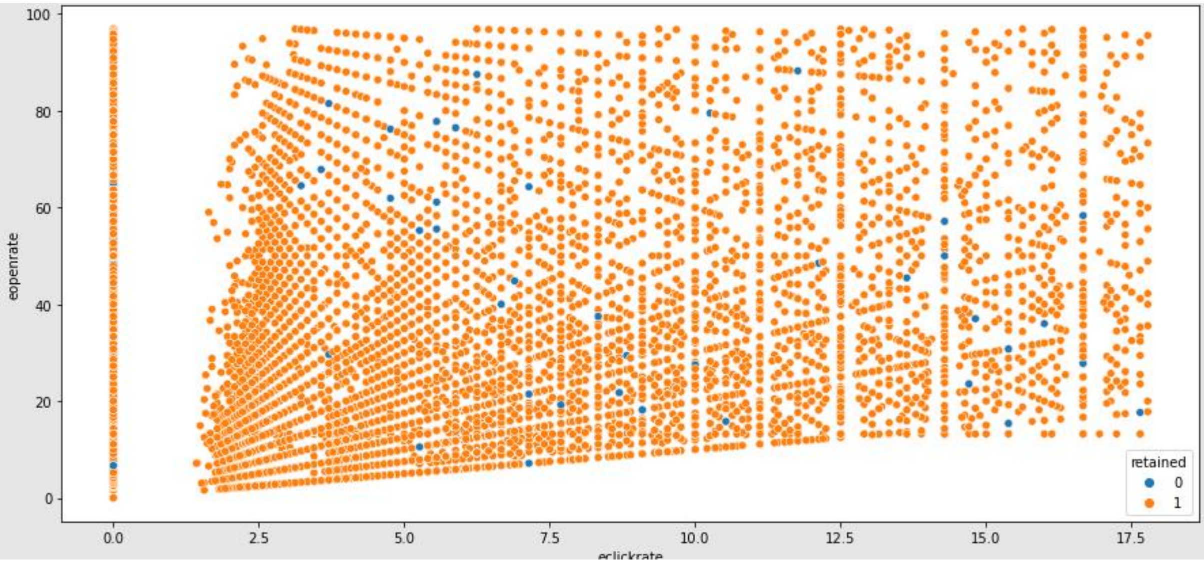
We have plotted **a scatter plot** for the fields **eclickrate** and **esent** to understand the retention trend.



Inferences -

- The greater the emails sent(esent) the eclickrate goes down.

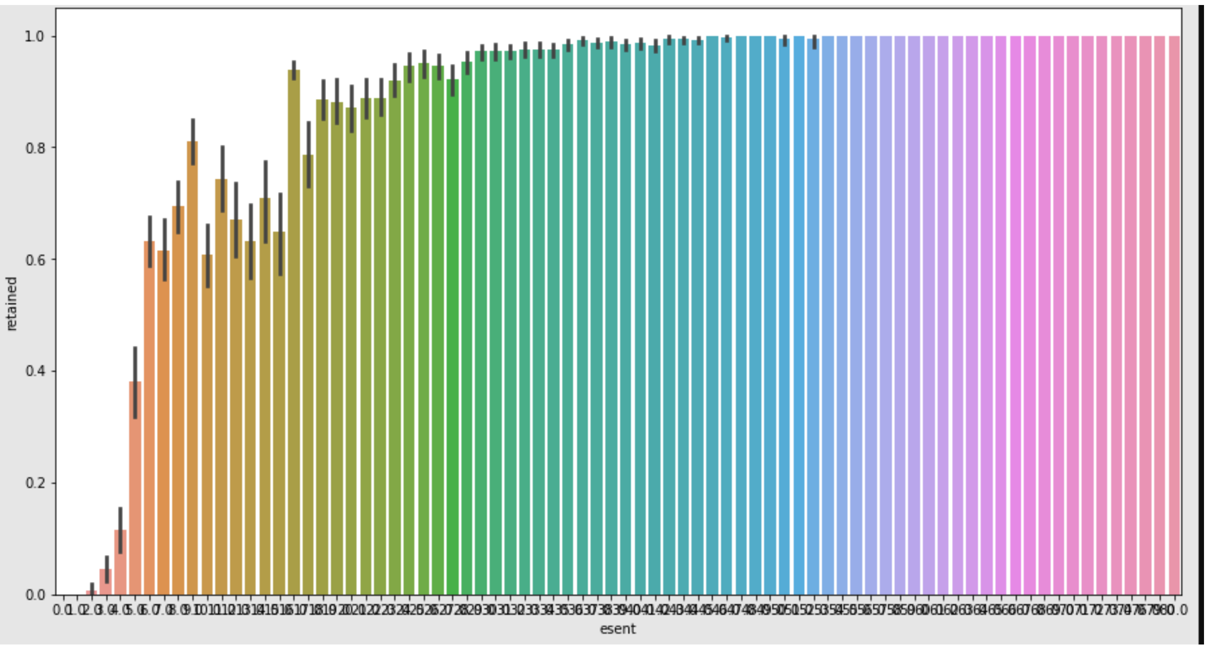
We have plotted a **scatter plot** for the fields **eclickrate** and **eopenrate** to understand the retention trend.



Inferences -

- As esent increases customers are likely to be retained

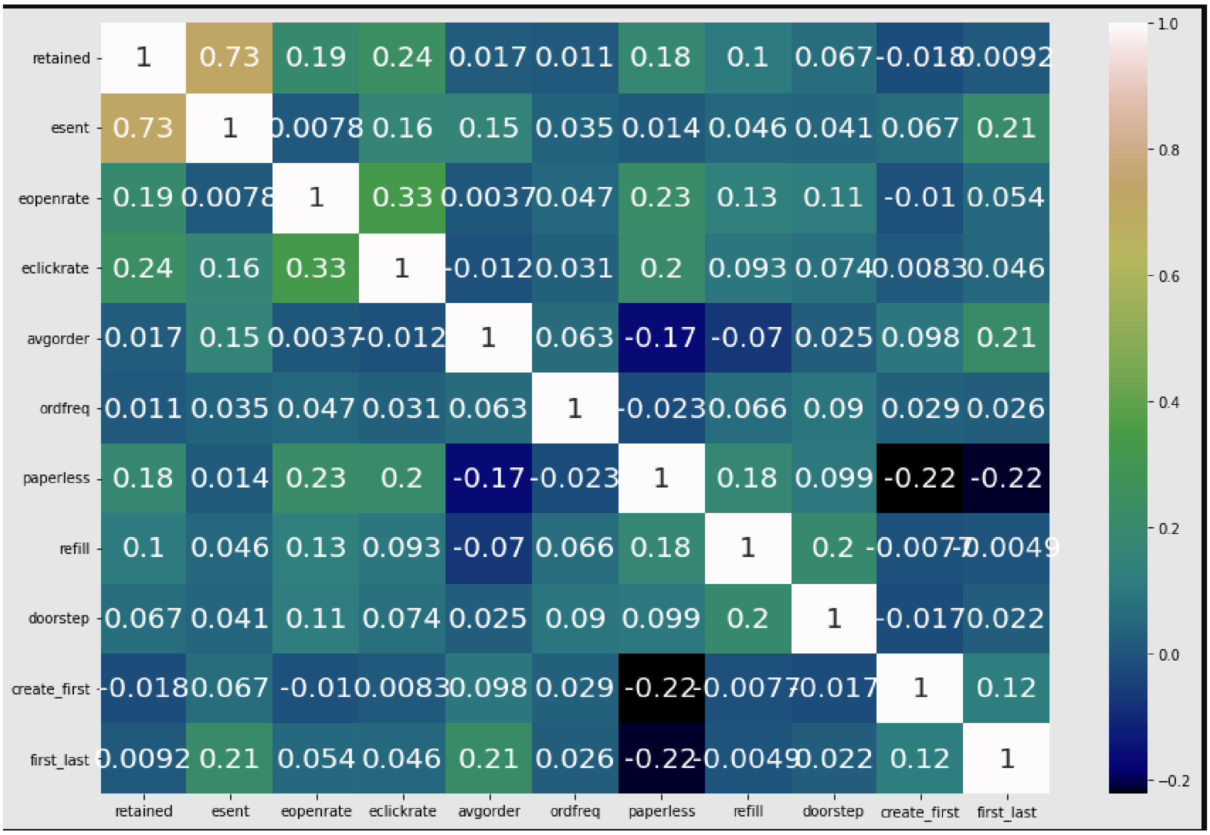
We have plotted **a bar plot** for the fields **esent** and **retained** to understand the retention trend.



Inferences -

- As esent increases customers are likely to be retained

We plotted a heatmap to check the correlation



Inferences -

- Positive correlation for esent and retained
- Positive correlation(moderate) for esent and eopenrate
- Paperless and Avg order are negatively correlated

Feature Engineering:

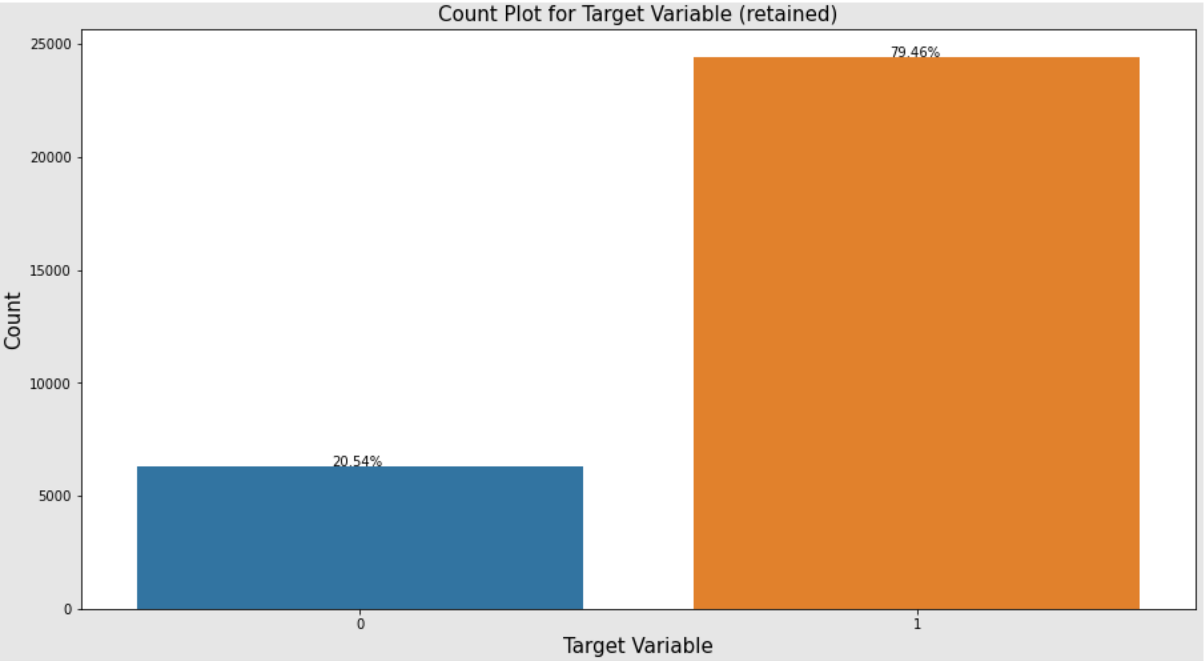
We have used transformation techniques like boxcox and we have scaled the data using standard scaler method.

Transformation techniques -

We have observed that few of our columns such as eopen rate,eclickrate ,avgorder and order frequency are highly right skewed hence we are using boxcox to normalize the data.

We have used the <u>standard scalar</u> to scale the data.Scaling makes it easy for the model to learn and understand the problem.

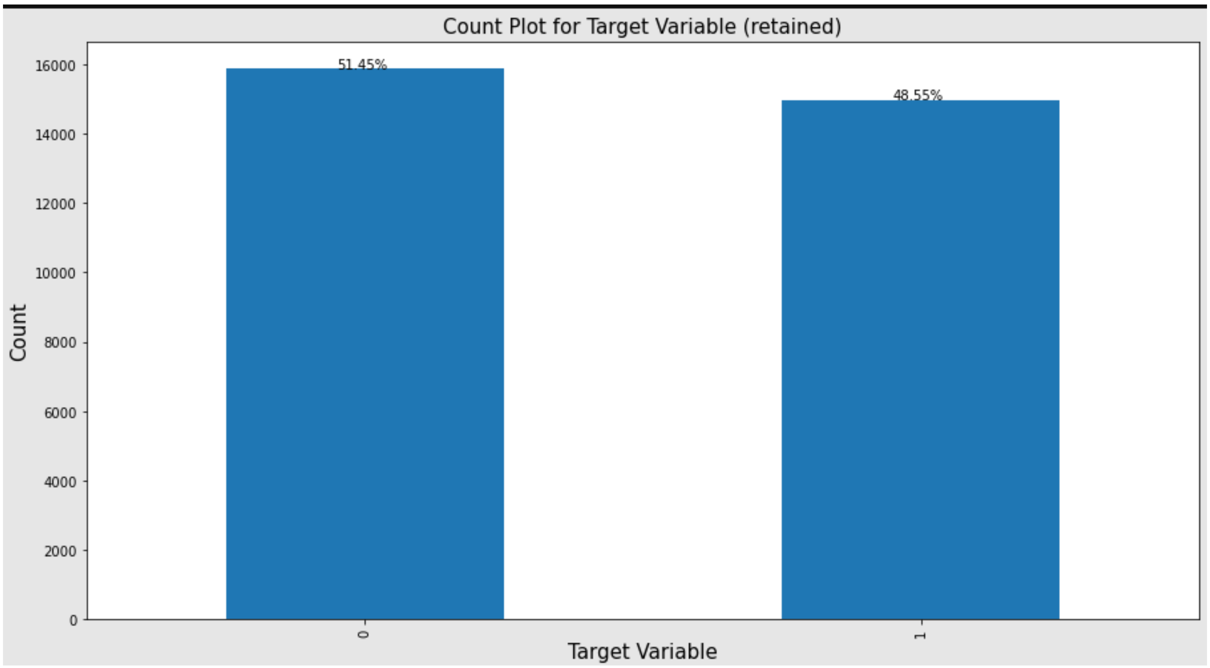
Class imbalance and it's treatment:



We can infer that the data is imbalanced. To treat it we can use

Imbalanced classification involves developing predictive models on classification datasets that have a severe class imbalance.
The challenge of working with imbalanced datasets is that most machine learning techniques will ignore, and in turn have poor performance on, the minority class, although typically it is performance on the minority class that is most important.
One approach to addressing imbalanced datasets is to oversample the minority class. The simplest approach involves duplicating examples in the minority class, although these examples don't add any new information to the model. Instead, new examples can be synthesized from the existing examples. This is a type of data augmentation for the minority class and is referred to as the Synthetic Minority Oversampling Technique, or SMOTE for short.

After Treating the imbalanced data, we can now observe that the data is balanced.



Base Model

We have used logistic regression and we have fitted out training data. We have predicted the ypred and we have used list comprehension by assuming that threshold value as (>0.5 - Retained & <0.5 - Not Retained).

We have plotted a confusion matrix to check whether our model has predicted the churn rate.

By using the classification report we can find the churn precision, churn recall and Churn f1 score.

To find the accuracy score and Roc_Auc_score we are using a function from sklearn metrics.

	Churn_precision	Churn_recall	Churn_f1score	Accuracy	Roc_Auc_Score
Base_model	0.70	0.90	0.79	0.904	0.904

VIF (Variance Inflation Factor)

To check Multi collinearity

	Features	VIF
5	paperless	2.760204
16	city_BOM	2.713546
17	city_DEL	2.555723
18	city_MAA	2.511845
10	favday_Monday	1.901113
14	favday_Tuesday	1.831598
0	esent	1.781733
1	eopenrate	1.718285
15	favday_Wednesday	1.666793
2	eclickrate	1.664925
13	favday_Thursday	1.635623
9	first_last	1.317455
4	ordfreq	1.310845
12	favday_Sunday	1.213729
6	refill	1.194474
8	create_first	1.161571
7	doorstep	1.132647
11	favday_Saturday	1.118166
3	avgorder	1.117228

As the value of every feature is less than the 5 we can say that the presence of multicollinearity is not present

Feature Selection

RFE is popular because it is easy to configure and use.It is effective in selecting those features in a training dataset that are most relevant in predicting the target variable.

```
Optimization terminated successfully.
         Current function value: 0.173057
         Iterations 8
                           Logit Regression Results
==============================================================================
Dep. Variable:               retained   No. Observations:                30895
Model:                          Logit   Df Residuals:                    30886
Method:                           MLE   Df Model:                            8
Date:                Fri, 18 Nov 2022   Pseudo R-squ.:                  0.7502
Time:                        20:19:51   Log-Likelihood:                -5346.6
converged:                       True   LL-Null:                       -21401.
Covariance Type:            nonrobust   LLR p-value:                     0.000
===================================================================================
                      coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
esent               3.7050      0.046     81.267      0.000       3.616       3.794
ordfreq             5.6659      0.326     17.380      0.000       5.027       6.305
refill              2.3506      0.116     20.191      0.000       2.122       2.579
doorstep            2.7554      0.189     14.557      0.000       2.384       3.126
create_first       -0.5596      0.029    -19.547      0.000      -0.616      -0.503
first_last         -0.8910      0.033    -27.408      0.000      -0.955      -0.827
favday_Saturday     1.8592      0.150     12.375      0.000       1.565       2.154
favday_Sunday       2.2947      0.205     11.197      0.000       1.893       2.696
city_DEL            1.9863      0.062     32.097      0.000       1.865       2.108
===================================================================================
```

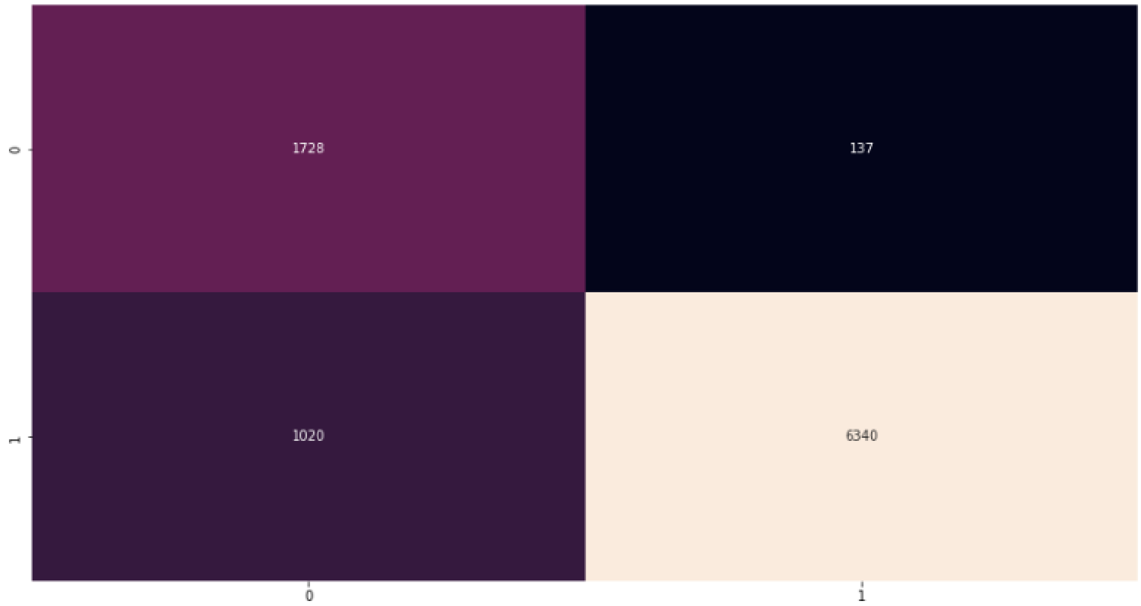
Using the Feature Selection RFE Model the features which are affecting the target variable are the following

```
0               esent
4             ordfreq
6              refill
7            doorstep
8        create_first
9          first_last
11    favday_Saturday
12      favday_Sunday
17           city_DEL
Name: Feature, dtype: object
```

Confusion Matrix for RFE model :

We have plotted a confusion matrix to check whether our model has predicted the churn rate.

```
Confusion Matrix: [[1728  137]
 [1020 6340]]
```

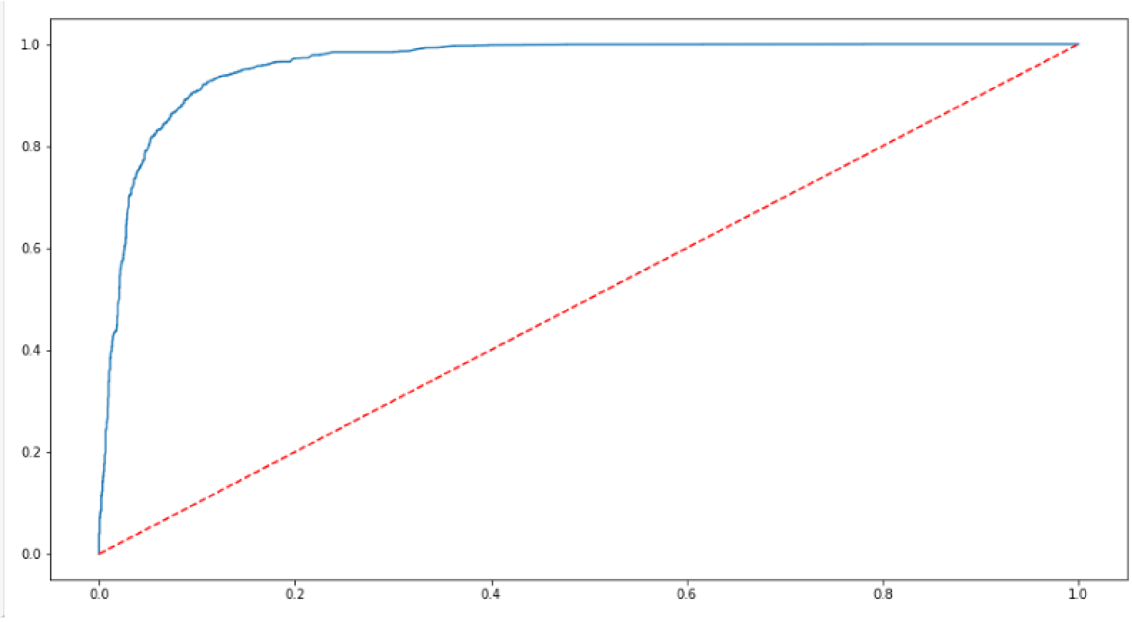


By using the classification report we can find the churn precision, churn recall and Churn f1 score.

To find the accuracy score and Roc_Auc_score we are using a function from sklearn metrics.

```
RFE Feature selected Logistic Regression:
              precision    recall  f1-score   support

           0       0.63      0.93      0.75      1865
           1       0.98      0.86      0.92      7360

    accuracy                           0.87      9225
   macro avg       0.80      0.89      0.83      9225
weighted avg       0.91      0.87      0.88      9225

Cohen Kappa Score: 0.67
ROC AUC Score: 0.894
```

ROC AUC Curve for RFE model



Hyper parameter tuning :

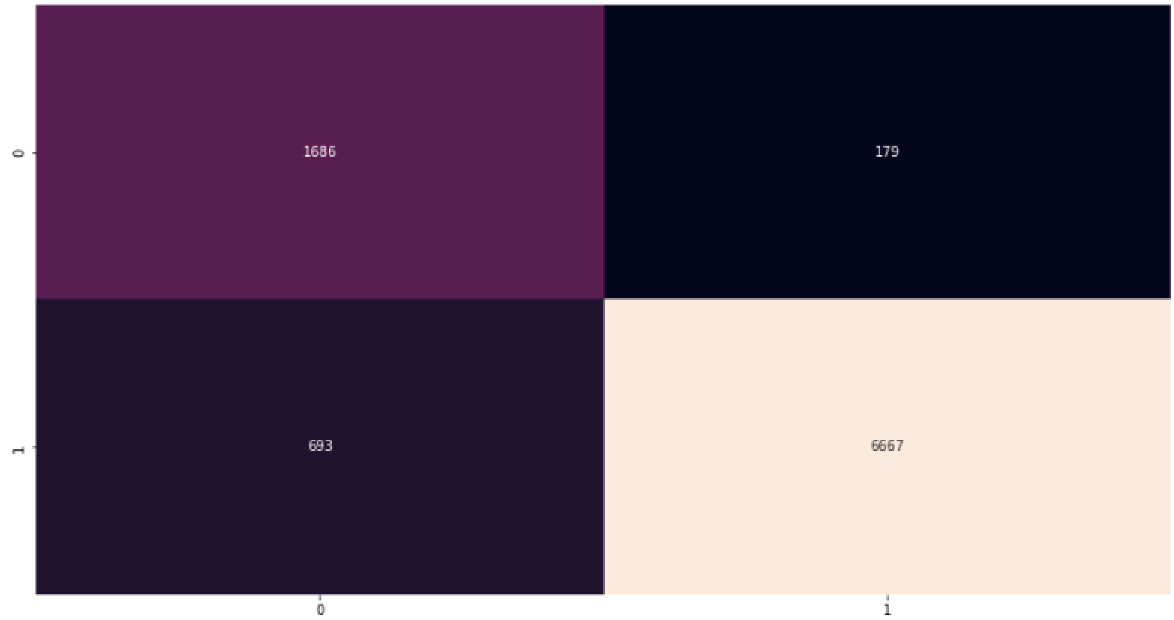
Decision Tree

In Decision Tree we are using hyper parameter tuning using GridSearchCV & we are fitting the training data to get the best parameters.

We have tuned the model using those parameters and fitted the model.Post that we have predicted the ypred.

We have plotted a confusion matrix to check whether our model has predicted the churn rate.

```
Confusion Matrix: [[1686  179]
 [ 693 6667]]
```

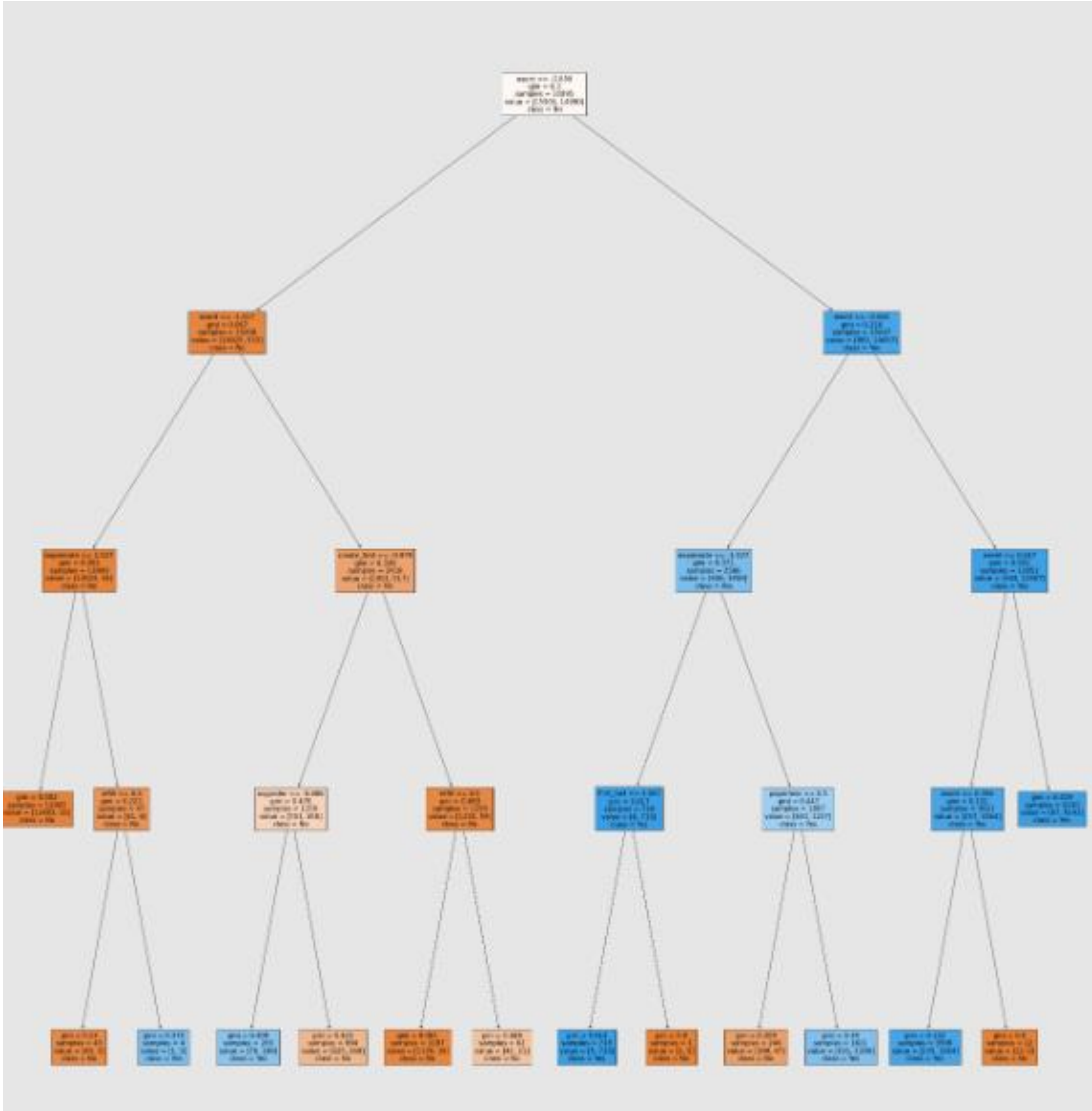


By using the classification report we can find the churn precision, churn recall and Churn f1 score.

To find the accuracy score and Roc_Auc_score we are using a function from sklearn metrics.

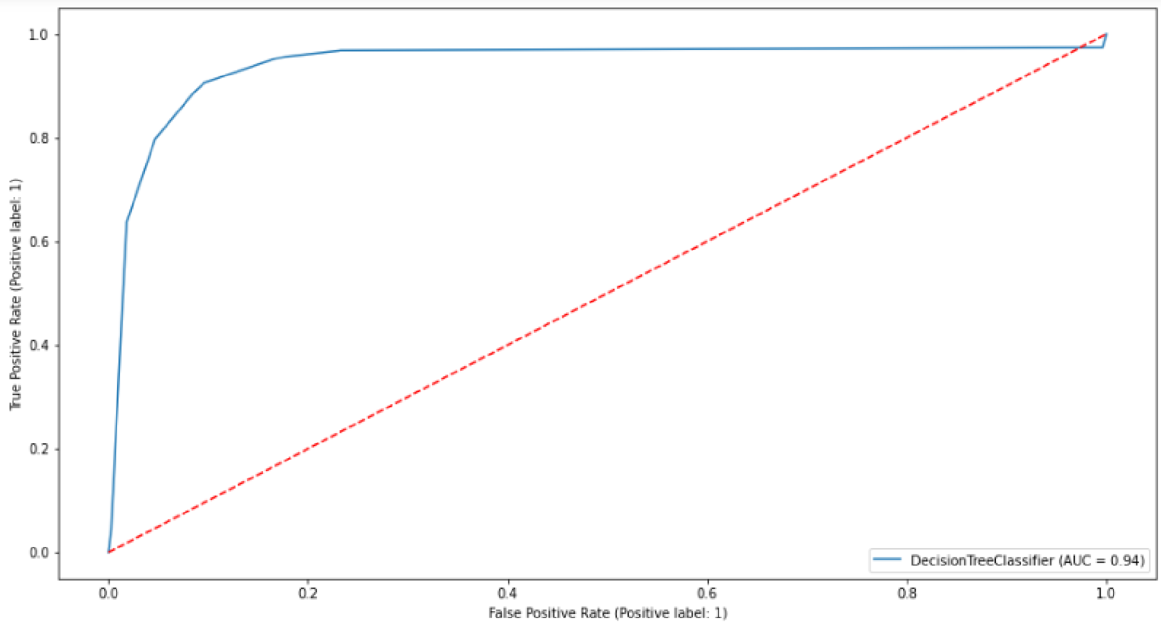
```
Decision Tree:
              precision    recall  f1-score   support

           0       0.71      0.90      0.79      1865
           1       0.97      0.91      0.94      7360

    accuracy                           0.91      9225
   macro avg       0.84      0.90      0.87      9225
weighted avg       0.92      0.91      0.91      9225

Cohen Kappa Score: 0.734
ROC AUC Score: 0.94
```

We have plotted the decision tree



ROC AUC Curve for decision tree

ROC_AUC score for decision tree is 0.905

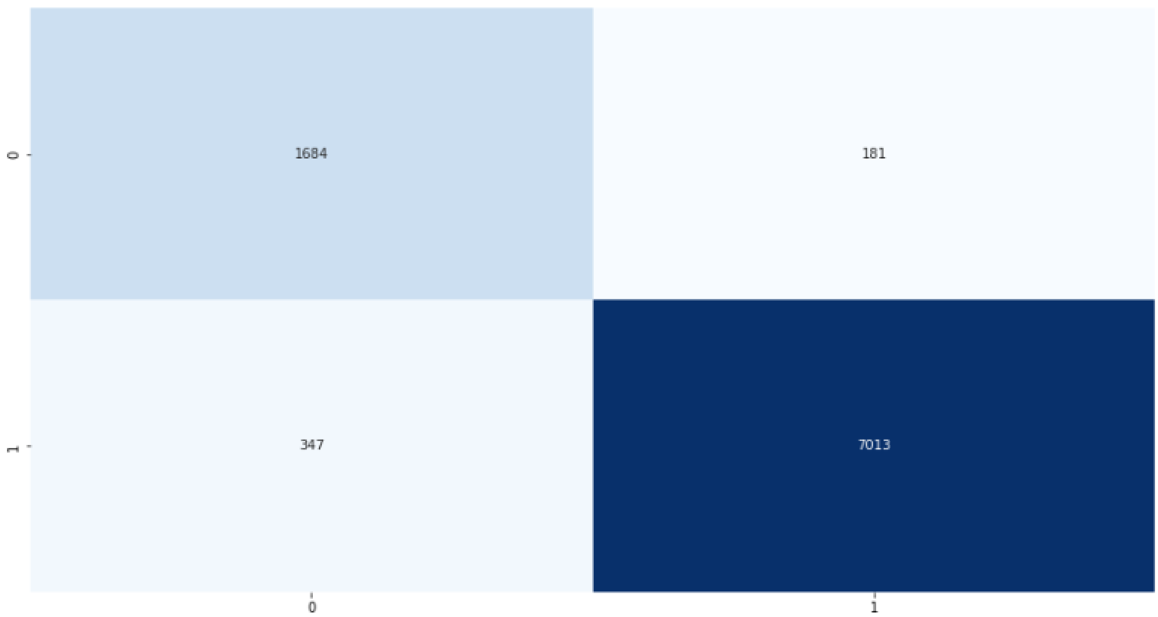


Random Forest

In Random Forest we are using hyper parameter tuning using GridSearchCV & we are fitting the training data to get the best parameters.

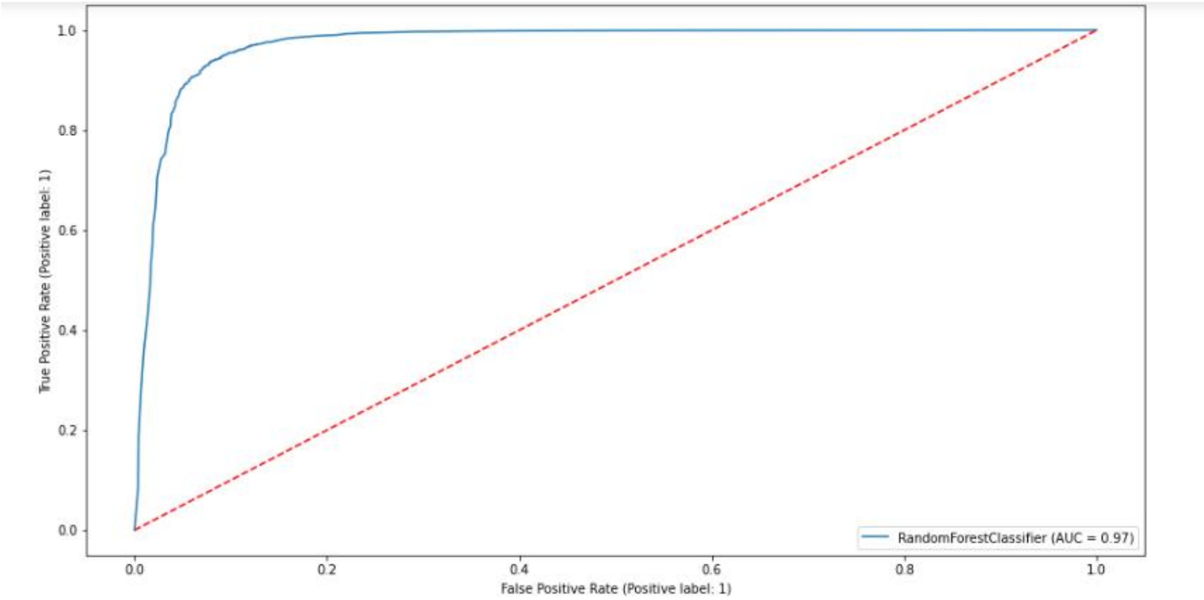
We have tuned the model using those parameters and fitted the model.Post that we have predicted the ypred.

We have plotted a confusion matrix to check whether our model has predicted the churn rate.



ROC AUC Curve for Random Forest

ROC_AUC score for Random Forest is 0.928



Boosting Techniques :

We have used two boosting methods.

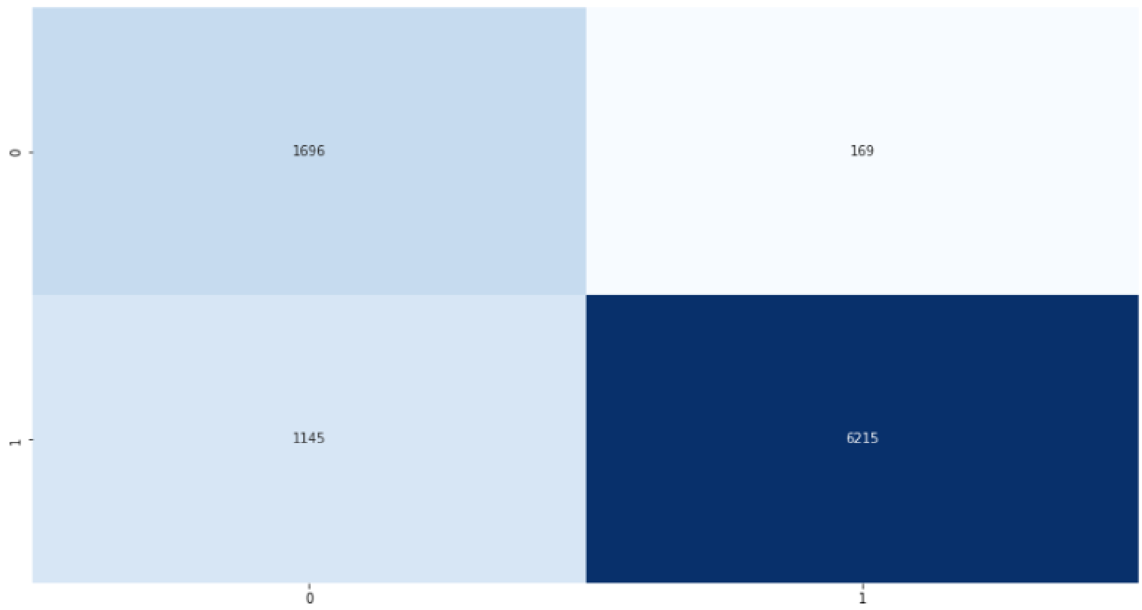
1.Adaboost

2.XGboost

Adaboost

In Adaboost we are using base estimator as decision tree classifier and fitting the X train,Y train and finding the ypred.

We have plotted a confusion matrix to check whether our model has predicted the churn rate.



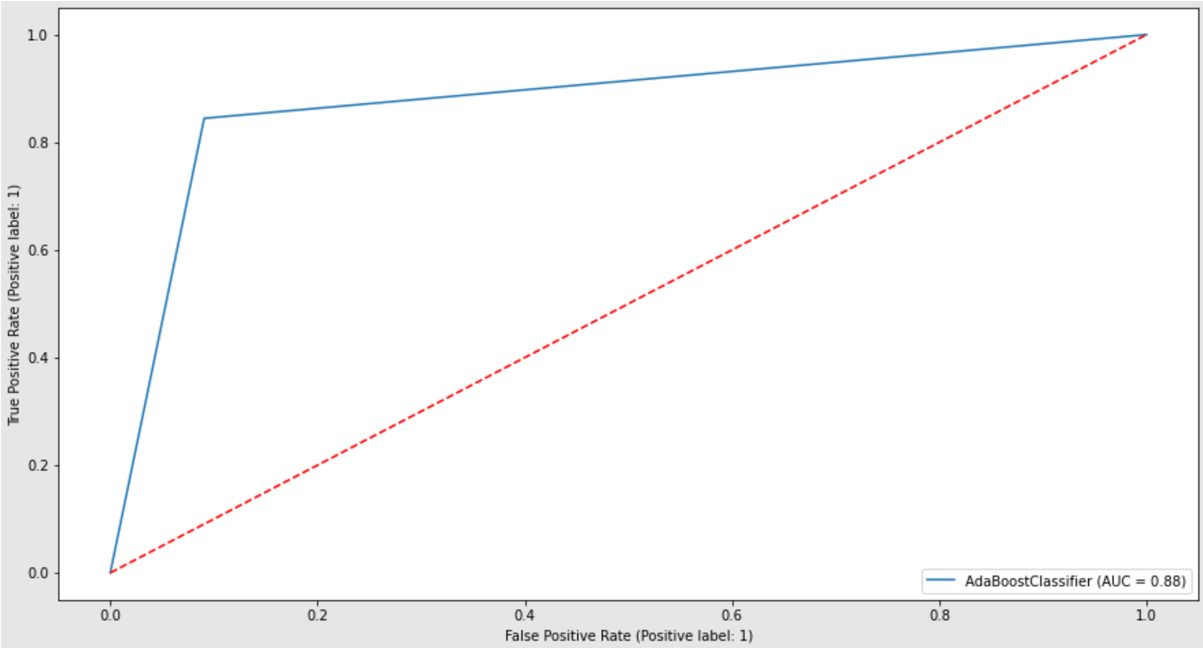
By using the classification report we can find the churn precision, churn recall and Churn f1 score.

To find the accuracy score and Roc_Auc_score we are using a function from sklearn metrics.

```
Adaboost:
              precision    recall  f1-score   support

           0       0.60      0.91      0.72      1865
           1       0.97      0.84      0.90      7360

    accuracy                           0.86      9225
   macro avg       0.79      0.88      0.81      9225
weighted avg       0.90      0.86      0.87      9225

Cohen Kappa Score: 0.63
```

ROC AUC Curve for Adaboost

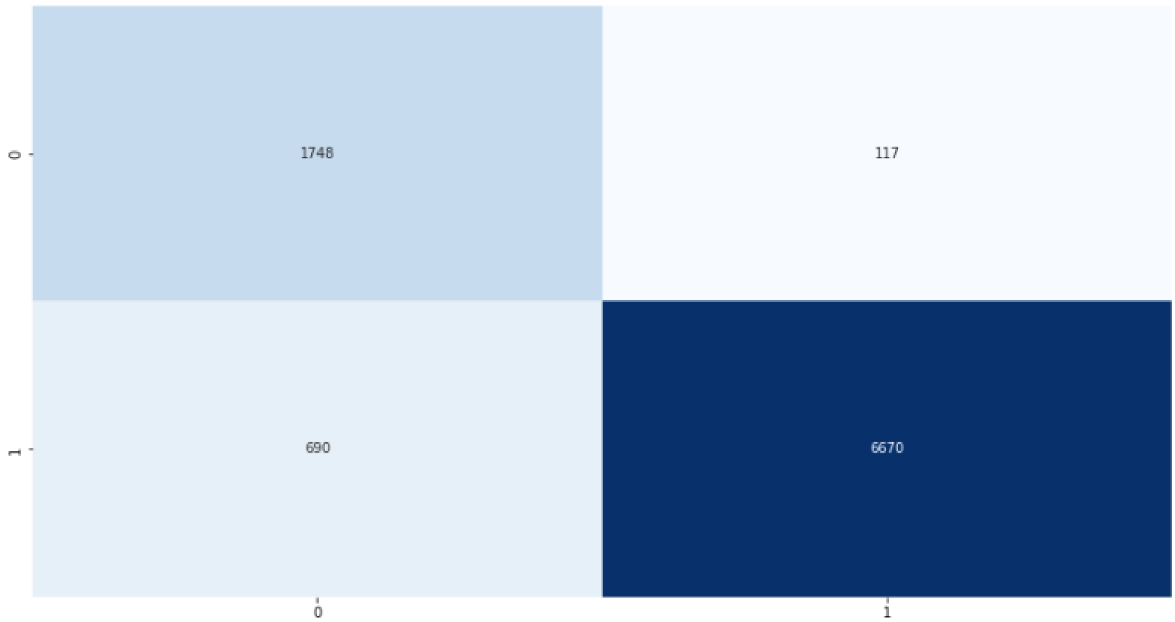
ROC_AUC score for Adaboost is 0.877



XGboost

In XGboost we are initiating the XGF classifier AND fitting the X train,Y train.Post that we find the ypred.

We have plotted a confusion matrix to check whether our model has predicted the churn rate.



By using the classification report we can find the churn precision, churn recall and Churn f1 score.

To find the accuracy score and Roc_Auc_score we are using a function from sklearn metrics.

```
XG Boost
              precision    recall  f1-score   support

           0       0.72      0.94      0.81      1865
           1       0.98      0.91      0.94      7360

    accuracy                           0.91      9225
   macro avg       0.85      0.92      0.88      9225
weighted avg       0.93      0.91      0.92      9225

COhen Kappa Score: 0.76
```

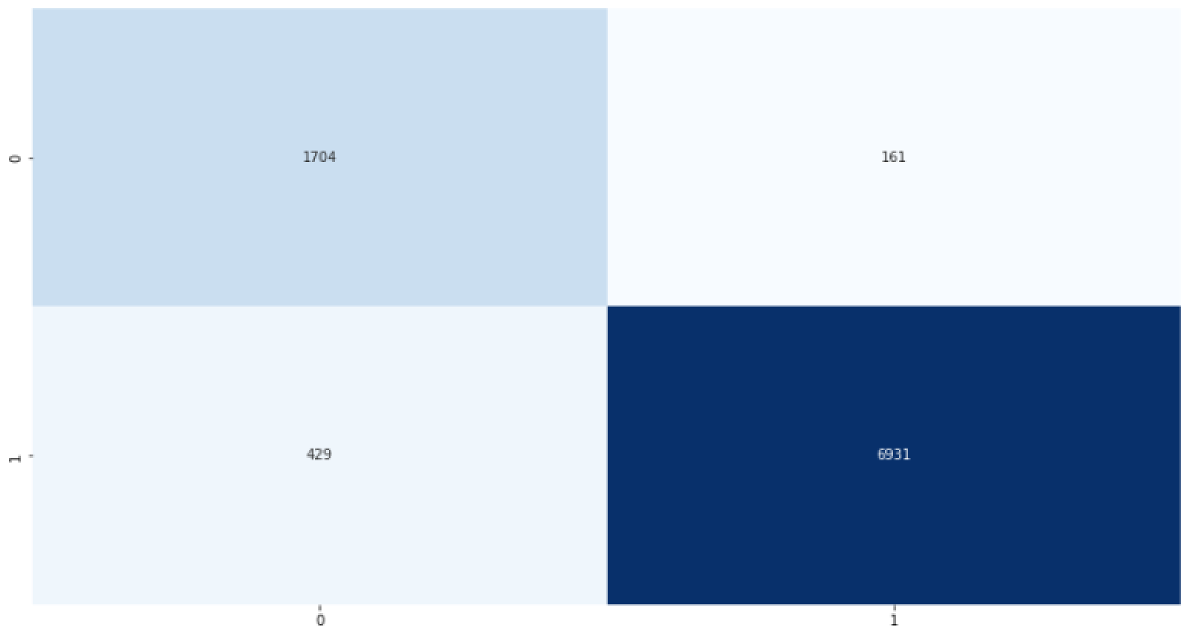
ROC AUC Curve for XGboost

ROC_AUC score for XGboost is 0.92

Stacking

In stacking we have used the base as decision tree classifier and random forest classified and Ada boost classifier as the final estimator. Post that we find the ypred.

We have plotted a confusion matrix to check whether our model has predicted the churn rate.



By using the classification report we can find the churn precision, churn recall and Churn f1 score.

To find the accuracy score and Roc_Auc_score we are using a function from sklearn metrics.

```
Stacking:
              precision    recall  f1-score   support

           0       0.80      0.91      0.85      1865
           1       0.98      0.94      0.96      7360

    accuracy                           0.94      9225
   macro avg       0.89      0.93      0.91      9225
weighted avg       0.94      0.94      0.94      9225

Cohen Kappa Score 0.81
```

ROC AUC Curve for Stacking

ROC_AUC score for Stacking is 0.928

Metric Table

	Churn_precision	Churn_recall	Churn_f1score	Accuracy	Roc_Auc_Score
Base_model	0.70	0.90	0.79	0.902000	0.903000
Rfe_Features_model	0.63	0.93	0.75	0.875000	0.894000
Decision_Tree	0.71	0.90	0.79	0.905000	0.905000
Random_Forest	0.83	0.90	0.86	0.943000	0.928000
AdaBoost	0.60	0.91	0.72	0.858000	0.877000
XGBoost	0.72	0.94	0.81	0.913000	0.922000
Stacking_Reg	0.80	0.91	0.85	0.936000	0.928000

The accuracy of the Random_Forest model is comparitively better than other models.Accuracy of Random Forest Model - 94.3% Hence we have selected the Random_Forest model.

Comparison & Selection of Model :

1. The accuracy of the Random_Forest model is comparitively better than other models.Accuracy of Random Forest Model - 94.3% The Stacking model gave a similar accuracy of 93.6% but since the random forest had a better presicion rate and f1 score,we have selected the Random_Forest model.

Results and Discussion :

1. Based on the results it is evident that our dataset performs better with the Random_Forest Model.
2. The accuracy of the Random_Forest model is comparitively better than other models.Accuracy of Random Forest Model - 94.3%
3. Features affecting our target variable - esent,favday_Sunday,favday_Saturday,city_DEL,create_first,doorstep,first_last,ordfreq,refill

Limitations :

1. Large investment in terms of price and time: Customer retention can prove expensive for business in the way that it involves large investment both in terms of price and time.
2. It requires huge cost for running loyalty programs in order to retain customers for a longer period. Business need to sacrifice their profit by offering several discount and cashback offers to its audience.
3. Require concerted commitment and Business Culture: Every business organization for attaining better customer retention rate should ensure a concerted commitment and proper culture. Every member working at distinct hierarchy of organization should focuses on providing best services to their customers.
4. New customers may be overlooked: Organization making efforts for attaining efficient customer retention rate may not focus on needs of new customers. There may be chances of new clients being overlooked by brand in a hoard to satisfy its existing customers. Unsatisfied customers may spread negative piece of information about the brand in society.
5. The dataset includes data from the year 2008-2018.The First five years 2008-2013 the customer behaviour towards online shopping would have been different as ecommerce had not boomed by then.
6. The Customer behaviour can be unpredictable.

Conclusion :

1. Targetted Email campaign for the existing customers will increase customer retention rate.
2. Faster 2-day delivery is an effective way to increase customer retention rate.
3. By using the Random forest model we can predict the churn rate with an accuracy of 94.3%

Recommendations :

1. Strong Customer service strategy to create a memorable experience to ensure repeat purchase.
2. Be a part of the social conversation.
3. Give gifts, rewards and prizes that are personalized and relevant to their interests
4. Create exclusive offers, hold special events, grant them early access to new offerings and VIP services.
5. Involve them in your new product planning processes.
6. Feature them in case studies and testimonials and podcasts.
7. Create a loyalty program

Assumptions to be satisfied :

- The logistic regression assumes that there is minimal or no multicollinearity among the independent variables.
- The Logistic regression assumes that the independent variables are linearly related to the log of odds.
- Logistic regression usually requires a large sample size to predict properly.
- Logistic regression assumes that the dependent variable is binary
- The Logistic regression assumes the observations to be independent of each other.
- We have assumed that the threshold value to be 0.5 in the base model.

References :

1. Ascarza, E. (2018). Retention futility: Targeting high-risk customers might be ineffective. Journal of Marketing Research, 55(1), 80–98. https://doi.org/10.1509/jmr.16.0163
2. Abu Zayyad, H.M., Obeidat, Z.M., Alshurideh, M.T., Abuhashesh, M., Maqableh, M., Masa'deh, R.E.: Corporate social responsibility and patronage intentions: the mediating effect of brand credibility. J. Mark. Commun. 1–24 (2020)
3. Alshurideh, D.M.: Do electronic loyalty programs still drive customer choice and repeat purchase behaviour? Int. J. Electron. Cust. Relatsh. Manag. 12(1), 40–57 (2019)
4. Zeithaml, V.A., Berry, L.L., Parasuraman, A.: The behavioral consequences of service quality. J. Mark. 60(2), 31–46 (1996)
5. Danesh, S.N., Nasab, S.A., Ling, K.C.: The study of customer satisfaction, customer trust and switching barriers on customer retention in Malaysia hypermarkets. Int. J. Bus. Manag. 7 (7), 141–150 (2012)
6. Al-Dmour, H., Alshuraideh, M., Salehih, S.: A study of Jordanians' television viewers habits. Life Sci. J. 11(6), 161–171 (2014)
7. Alshurideh, M.: The factors predicting students' satisfaction with universities' healthcare clinics' services: a case-study from the Jordanian higher education sector. Dirasat. Adm. Sci. 161(1524), 1–36 (2014)
8. Preikschas, M.W., Cabanelas, P., Rüdiger, K., Lampón, J.F.: Value co-creation, dynamic capabilities and customer retention in industrial markets. J. Bus. Ind. Mark. 32(3), 409–420 (2017)
9. Al Dmour, H., Alshurideh, M., Shishan, F.: The influence of mobile application quality and attributes on the continuance intention of mobile shopping. Life Sci. J. 11(10), 172–181 (2014)
10. Obeidat, B., Sweis, R., Zyod, D., Alshurideh, M.: The effect of perceived service quality on customer loyalty in internet service providers in Jordan. J. Manag. Res. 4(4), 224–242 (2012)
11. Al-dweeri, R., Obeidat, Z., Al-dwiry, M., Alshurideh, M., Alhorani, A.: The impact of eservice quality and e-loyalty on online shopping: moderating effect of e-satisfaction and etrust. Int. J. Mark. Stud. 9(2), 92–103 (2017)

12. Levesque, T., McDougall, G.H.G.: Determinants of customer satisfaction in retail banking. Int. J. bank Mark. 14(7), 12–20 (1996)
13. Hansemark, O.C., Albinsson, M.: Customer satisfaction and retention: the experiences of individual employees. Manag. Serv. Qual. Int. J. 14(1), 40–57 (2004)
14. Morgan, R.M., Hunt, S.D.: The commitment-trust theory of relationship marketing. J. Mark. 58(3), 20–38 (1994)