A PROJECT REPORT

on

"CLIENT ATTRITION FORECASTING"

Submitted to

KIIT Deemed to be University

In Partial Fulfilment of the Requirement for the Award of

BACHELOR'S DEGREE IN
COMPUTER SCIENCE AND ENGINEERING

BY

ADITYA KUMAR LAL	1605170
AMAN PATNAIK	1605174
ASHUTOSH SHARMA	1605184
SUMIT KUMAR SONI	1605241

UNDER THE GUIDANCE OF
PROF. TANMAYA SWAIN

SCHOOL OF COMPUTER ENGINEERING
KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY
BHUBANESWAR, ODISHA - 751024

April 2020

A PROJECT REPORT

on

"CLIENT ATTRITION FORECASTING"

Submitted to

KIIT Deemed to be University

In Partial Fulfilment of the Requirement for the Award of

BACHELOR'S DEGREE IN
COMPUTER SCIENCE AND ENGINEERING

BY

ADITYA KUMAR LAL	1605170
AMAN PATNAIK	1605174
ASHUTOSH SHARMA	1605184
SUMIT KUMAR SONI	1605241

UNDER THE GUIDANCE OF
PROF. TANMAYA SWAIN

SCHOOL OF COMPUTER ENGINEERING
KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY
BHUBANESWAE, ODISHA -751024

April 2020

KIIT Deemed to be University

School of Computer Engineering
Bhubaneswar, ODISHA 751024

CERTIFICATE

This is certify that the project entitled

"CLIENT ATTRITION FORECASTING"

submitted by

ADITYA KUMAR LAL	1605170
AMAN PATNAIK	1605174
ASHUTOSH SHARMA	1605184
SUMIT KUMAR SONI	1605241

is a record of bonafide work carried out by them, in the partial fulfilment of the requirement for the award of Degree of Bachelor of Engineering (Computer Science & Engineering) at KIIT Deemed to be university, Bhubaneswar. This work is done during year 2019-2020, under our guidance.

Date: /   /

(Prof. Co-Guide Name)
Project Co-Guide

(Prof. Tanmaya Swain)
Project Guide

Acknowledgements

We are profoundly grateful to Prof. TANMAYA SWAIN for his expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion.

ADITYA KUMAR LAL
AMAN PATNAIK
ASHUTOSH SHARMA
SUMIT KUMAR SONI

ABSTRACT

In this competitive world, business is becoming highly saturated and client becomes the lifeline of every business. The fate of the business is dependent on them. They are the primary growth pillars of every company. Moreover it is very expensive to get a new client. Hence the companies are trying to retain their existing clients. Now the companies are trying to know in advance the client likely to churn and how to retain those with some special attrition programs and perks.

Client attrition is a tendency of clients to leave a company and stop being a paying client of it. Companies with high attrition rates aren't only failing to deliver in their relationships with ex-clients but also damage their future acquisition efforts by creating negative word-of-mouth around their products. Companies that constantly monitor how client engage with products, encourage clients to share opinions, and solve their issues immediately have greater opportunities to maintain mutually beneficial client relationships.

This project explores the application of data mining techniques in forecasting the likely churners and attribute selection on identifying the churn. It also compares the efficiency of several classifiers and lists their performances for the datasets.

Keywords: data mining, attribute selection, attrition forecasting, machine learning.

Contents

SERIAL NUMBER	CONTENTS	PAGE NUMBER
1	Introduction	1
2	Software Requirements Specification	2
3	Requirement Analysis and Process Flow	3
4	Model Building and Data Analysis	4
5	Implementation	6
6	Screenshots of Project and Findings	9
7	Conclusion and Future Scope	15
8	References	16
9	Student's Contribution to the Project	17

List of Figures

FIGURE NUMBER	IMAGE CAPTION	PAGE NUMBER
3.1	Steps Followed	3
4.1	Process Flow	4
4.2	Process Flow	5
4.3	Data set	5
6.1	Gender Distribution	9
6.2	% of Senior Citizen	9
6.3	% Customers with dependents and partners	10
6.4	% Customers by their tenure	10
6.5	% Customers by contract type	10
6.6	Tenure of client based on their contract type	11
6.7	Distribution of various services used by clients	11
6.8	Distribution of TotalCharges vs. MonthlyCharges	12
6.9	Churn Rate	12
6.10	weights of all the variables	13
6.11	weights of all the variables	14
6.12	Churn Rate vs. Gender	14

Chapter 1

Introduction

One of the telecom service providers is struggling with sales & attrition. They want to retain their clients, so we analyze the client's data and develop a model that will focus on client attrition programs. According to the SAS Institute report [1], the annual rate of client churn in telecommunications industry is currently at about 30% with an upward trend in correlation with the growth of the market. The broad definition of churn can be said as the actions of a client's service is terminated either by the client themselves or the service provider for violation of service agreements [2], [3]. However, the major and most often cause of churn is the client due to non-satisfaction in the service by a provider or due to more enhanced affordable service by other service-provider [4], [5].

The dataset includes information about:

1. Clients who left the service provider.
2. Services that each client has signed up for like- phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies.
3. Client account information – how long they've been a client, contract, payment method, paperless billing, monthly charges, and total charges.
4. Demographic info about clients – gender, age, and if they have partners and dependents.

Client attrition is a major problem in telecom industry as it is expensive to get the new client. Hence the telecom companies are trying to retain their existing clients. Now the companies are trying to know in advance the clients likely to churn and how to retain those with some special attrition programs and perks [6].

Client attrition may be due to followings:

1. Pricing Promotions
2. Service Quality
3. Clients Service
4. New Competitor entering to market with competitive pricing

Chapter 2

Software Requirements Specification

2.1 Software Tools and Technologies:-

2.1.1 Data Pr-processing: Python, Anaconda, Pandas
2.1.2 Data Preparation: Scikit Learn, Statistics
2.1.3 Data Visualization: Matplotlib, Seaborn
2.1.4 Algorithms: Logistic Regression, Random Forest, SVM, XG Boost

2.2 Hardware Requirements:-

2.2.1 RAM: 8 GB
2.2.2 I5 Processor

Chapter 3
Requirement Analysis and Process Flow

Requirement analysis tries those tasks that go into determining the needs or conditions to meet for a new or altered product or project, taking various requirements into the account like: stakeholders, analyzing, documenting, validating and managing software or system requirements. Requirement analysis plays a major role in each and every project, before going for designation of any type of project we have to find out the detailed requirement of the project and accordingly have to work on it.

Data preparation and feature engineering:-

1. Handling Missing Values.
2. Dropping redundant features.
3. Categorical features into numerical.
4. SMOTE (Synthetic Minority Over-sampling Technique).
5. Splitting the data-set into training and testing.

Steps/outline followed:-

1. Use Case/Business Problem understanding
2. Data Acquisition and Preprocessing
3. Feature selection and Engineering
4. Exploratory Data Analysis(Visualizations)
5. Model Building/Development and validations
6. Insights and actions required to retain clients

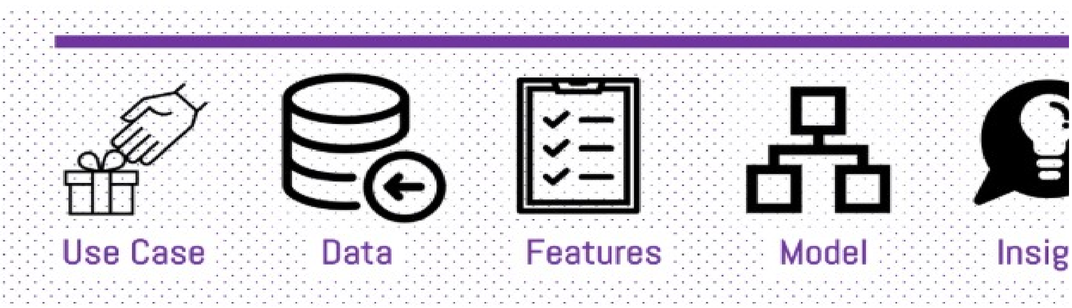

Figure 3.1: Steps Followed

Chapter 4

Model Building and Data Analysis

After requirement gathering and Analysis we proceed for the model building and raw data analysis, so that we can find the accurate data and evaluation can be made in order to get the required predicted result. Building a necessary model will give a total idea how to execute the total raw data and reduce the redundant data and we can able to predict the accurate data and get the exact predicted result and the corresponding model building process flow diagram given below.

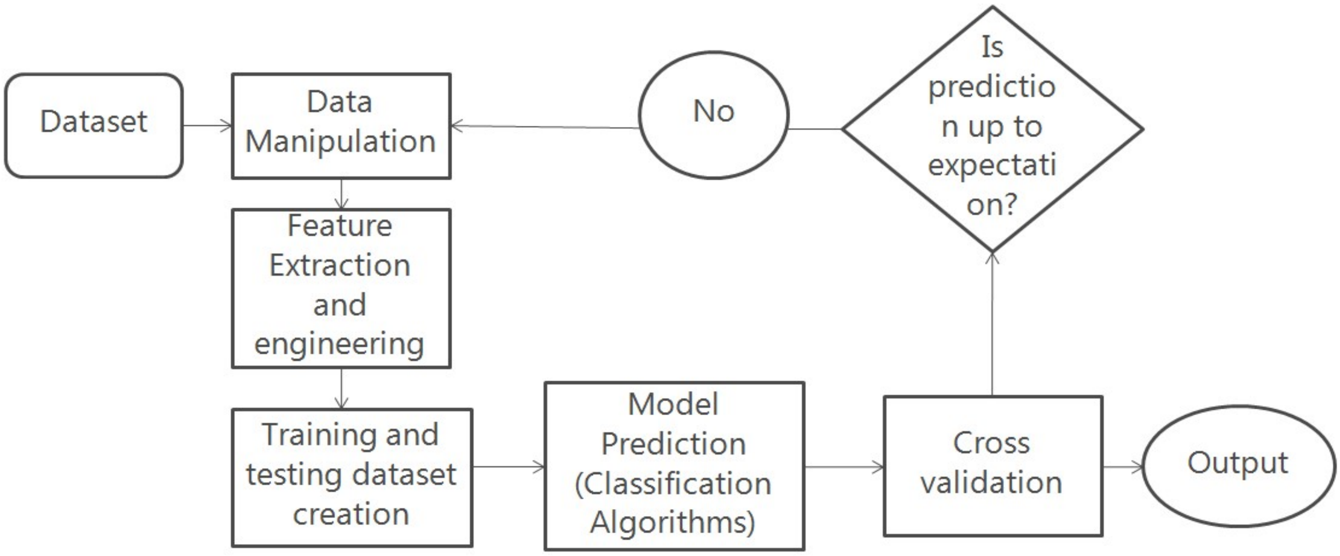

Figure 4.1: Process Flow

Data Analysis

The dataset contains 6970 clients' information with 21 features.

Client ID : Unique Id of each client

Gender : Gender of the client

Senior Citizen : Client whose age is above 60 years

Partner : Client is married/ in a live-in relationship

Dependents : Client has dependents (children/ retired parents)

Tenure : The time for which a client has been using the service.

PhoneService : Whether a client has a landline phone service along with the internet service.

MultipleLines : Whether a client has multiple lines of internet connectivity

InternetService : The type of internet services chosen by the client.

OnlineSecurity : Specifies if a client has online security.

OnlineBackup : Specifies if a client has online backup.

Device Protection : Specifies if a client has opted for device protection.

TechSupport : Whether a client has opted for tech support of not.

StreamingTV : Whether a client has an option of TV streaming.

StreamingMovies : Whether a client has an option of Movie streaming.

Contract : The type of contract a client has chosen.

DigitalBilling : Whether a client has opted for paperless billing.

PaymentMode : Specifies the method by which bills are paid.

MonthlyCharges : Specifies the money paid by a client each month.

TotalCharges : The total money paid by the client to the company.

Churn : Client has churned or not

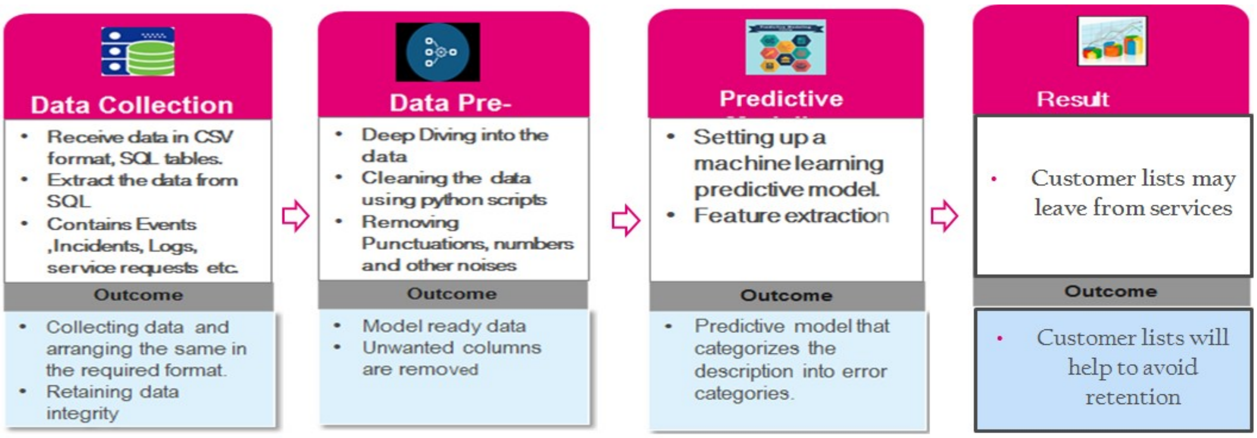

Figure 4.2: Process Flow

	customerID	Sex	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	Tec
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	

5 rows × 21 columns

Figure 4.3: Data set

Chapter 5
Implementation

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.ticker as mtick
import matplotlib.pyplot as plt
sns.set(style = 'white')
client = pd.read_csv("C:\\Users\\nEW u\\Desktop\\New folder\\Data_usecase2.csv")
client.head()
client.columns.values
client.dtypes
client.TotalCharges = pd.to_numeric(client.TotalCharges, errors='coerce')
client.isnull().sum()
client.dropna(inplace = True)
df2 = client.iloc[:,1:]
df2['Churn'].replace(to_replace='Yes', value=1, inplace=True)
df2['Churn'].replace(to_replace='No',  value=0, inplace=True)
df_dummies = pd.get_dummies(df2)
df_dummies.head()
colors = ['Orange','Green']
ax = (client['Gender'].value_counts()*100.0 /len(client)).plot(kind='bar',
                                                                stacked = True,
                                                               rot = 0,
                                                               color = colors)
ax.yaxis.set_major_formatter(mtick.PercentFormatter())
ax.set_ylabel('% Customers')
ax.set_xlabel('Gender')
ax.set_ylabel('% Customers')
ax.set_title('Gender Distribution')
totals = []
for i in ax.patches:
    totals.append(i.get_width())
total = sum(totals)
for i in ax.patches:
    ax.text(i.get_x()+.15, i.get_height()-3.5, \
            str(round((i.get_height()/total), 1))+'%',
            fontsize=12,
            color='white',
           weight = 'bold')
ax = (client['SeniorCitizen'].value_counts()*100.0 /len(client))\
.plot.pie(autopct='%.1f%%', labels = ['No', 'Yes'],figsize =(6,6), fontsize = 14)
ax.yaxis.set_major_formatter(mtick.PercentFormatter())
ax.set_ylabel('Senior Citizens',fontsize = 14)
ax.set_title('% of Senior Citizens', fontsize = 14)
df2 = pd.melt(client, id_vars=['customerID'], value_vars=['Dependents','Married'])
df3 = df2.groupby(['variable','value']).count().unstack()
df3 = df3*100/len(client)
colors = ['Orange','Green']
ax = df3.loc[:,'customerID'].plot.bar(stacked=True, color=colors,
                                      figsize=(8,6),rot = 0,
                                     width = 0.2)
ax.yaxis.set_major_formatter(mtick.PercentFormatter())
ax.set_ylabel('% Customers',size = 14)
ax.set_xlabel('')
ax.set_title('% Customers with dependents and partners',size = 14)
ax.legend(loc = 'center',prop={'size':14})
for p in ax.patches:
    width, height = p.get_width(), p.get_height()
    x, y = p.get_xy()
    ax.annotate('{:.0f}%'.format(height), (p.get_x()+.25*width, p.get_y()+.4*height),
                color = 'white',
               weight = 'bold',
               size = 14)
```

```
ax = sns.distplot(client['tenure'], hist=True, kde=False,
             bins=int(180/5), color = 'red',
             hist_kws={'edgecolor':'black'},
             kde_kws={'linewidth': 4})
ax.set_ylabel('# of Customers')
ax.set_xlabel('Tenure (months)')
ax.set_title('# of Customers by their tenure')
ax = client['Contract'].value_counts().plot(kind = 'bar',rot = 0, width = 0.3, color = 'red')
ax.set_ylabel('# of Customers')
ax.set_title('# of Customers by Contract Type')
fig, (ax1,ax2,ax3) = plt.subplots(nrows=1, ncols=3, sharey = True, figsize = (20,6))
ax = sns.distplot(client[client['Contract']=='Month-to-month']['tenure'],
                   hist=True, kde=False,
                   bins=int(180/5), color = 'Orange',
                   hist_kws={'edgecolor':'black'},
                   kde_kws={'linewidth': 4},
                 ax=ax1)
ax.set_ylabel('# of Customers')
ax.set_xlabel('Tenure (months)')
ax.set_title('Month to Month Contract')
ax = sns.distplot(client[client['Contract']=='One year']['tenure'],
                   hist=True, kde=False,
                   bins=int(180/5), color = 'white',
                   hist_kws={'edgecolor':'black'},
                   kde_kws={'linewidth': 4},
                 ax=ax2)
ax.set_xlabel('Tenure (months)',size = 14)
ax.set_title('One Year Contract',size = 14)
ax = sns.distplot(client[client['Contract']=='Two year']['tenure'],
                   hist=True, kde=False,
                   bins=int(180/5), color = 'Green',
                   hist_kws={'edgecolor':'black'},
                   kde_kws={'linewidth': 4},
                 ax=ax3)
ax.set_xlabel('Tenure (months)')
ax.set_title('Two Year Contract')
client.columns.values
services = ['PhoneService','MultipleLines','InternetService','OnlineSecurity',
           'OnlineBackup','DeviceProtection','TechSupport','StreamingTV','StreamingMovies']
fig, axes = plt.subplots(nrows = 3,ncols = 3,figsize = (15,12))
for i, item in enumerate(services):
    if i < 3:
        ax = client[item].value_counts().plot(kind = 'bar',ax=axes[i,0],rot = 0)
    elif i >=3 and i < 6:
        ax = client[item].value_counts().plot(kind = 'bar',ax=axes[i-3,1],rot = 0)
    elif i < 9:
        ax = client[item].value_counts().plot(kind = 'bar',ax=axes[i-6,2],rot = 0)
    ax.set_title(item)
client[['MonthlyCharges', 'TotalCharges']].plot.scatter(x = 'MonthlyCharges',
                                                          y='TotalCharges',color='orange')
colors = ['Orange','Green']
ax = (client['Churn'].value_counts()*100.0 /len(client)).plot(kind='bar',
                                                               stacked = True,
                                                              rot = 0,
                                                              color = colors,
                                                             figsize = (7,5))
ax.yaxis.set_major_formatter(mtick.PercentFormatter())
ax.set_ylabel('% Customers',size = 14)
ax.set_xlabel('Churn',size = 14)
ax.set_title('Churn Rate', size = 14)
totals = []
for i in ax.patches:
    totals.append(i.get_width())
total = sum(totals)
for i in ax.patches:
    # get_width pulls left or right; get_y pushes up or down
    ax.text(i.get_x()+.15, i.get_height()-4.0, \
            str(round((i.get_height()/total), 1))+'%',
            fontsize=12,
            color='white',
           weight = 'bold',
           size = 14)
```

```
sns.barplot(x = client.Churn, y = client.tenure,color= 'red')
ax = sns.kdeplot(client.MonthlyCharges[(client["Churn"] == 'No') ],
             color="Red", shade = True)
ax = sns.kdeplot(client.MonthlyCharges[(client["Churn"] == 'Yes') ],
             ax =ax, color="Blue", shade= True)
ax.legend(["Not Churn","Churn"],loc='upper right')
ax.set_ylabel('Density')
ax.set_xlabel('Monthly Charges')
ax.set_title('Distribution of monthly charges by churn')
ax = sns.kdeplot(client.TotalCharges[(client["Churn"] == 'No') ],
             color="Red", shade = True)
ax = sns.kdeplot(client.TotalCharges[(client["Churn"] == 'Yes') ],
             ax =ax, color="Blue", shade= True)
ax.legend(["Not Churn","Churn"],loc='upper right')
ax.set_ylabel('Density')
ax.set_xlabel('Total Charges')
ax.set_title('Distribution of total charges by churn')
y = df_dummies['Churn'].values
X = df_dummies.drop(columns = ['Churn'])
from sklearn.preprocessing import MinMaxScaler
features = X.columns.values
scaler = MinMaxScaler(feature_range = (0,1))
scaler.fit(X)
X = pd.DataFrame(scaler.transform(X))
X.columns = features
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=101)
from sklearn.linear_model import LogisticRegression
model = LogisticRegression()
result = model.fit(X_train, y_train)
from sklearn import metrics
prediction_test = model.predict(X_test)
print (metrics.accuracy_score(y_test, prediction_test))
weights = pd.Series(model.coef_[0],
              index=X.columns.values)
print (weights.sort_values(ascending = False)[:10].plot(kind='bar'))
print(weights.sort_values(ascending = False)[-10:].plot(kind='barh'))
from sklearn.ensemble import RandomForestClassifier
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=101)
model_rf = RandomForestClassifier(n_estimators=1000 , oob_score = True, n_jobs = -1,
                          random_state =50, max_features = "auto",
                          max_leaf_nodes = 30)
model_rf.fit(X_train, y_train)
prediction_test = model_rf.predict(X_test)
print (metrics.accuracy_score(y_test, prediction_test))
from sklearn.ensemble import AdaBoostClassifier
model = AdaBoostClassifier()
model.fit(X_train,y_train)
preds = model.predict(X_test)
metrics.accuracy_score(y_test, preds)
from xgboost import XGBClassifier
model = XGBClassifier()
model.fit(X_train, y_train)
preds = model.predict(X_test)
metrics.accuracy_score(y_test, preds)
```

Chapter 6
Screenshots of Project and Findings

6.1 Gender Distribution- About 50.9% of the clients in our dataset is male while the 49.1% are female.

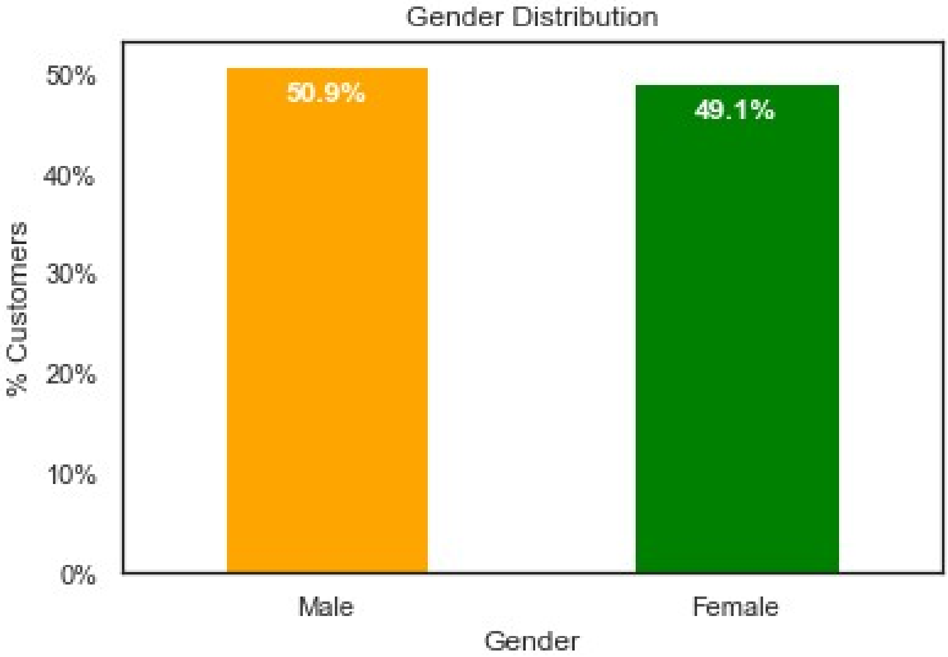

Figure 6.1: Gender Distribution

6.2 %Senior Citizens- There is only 16% of the clients who are senior citizens. Thus most of our clients in the data are young people.

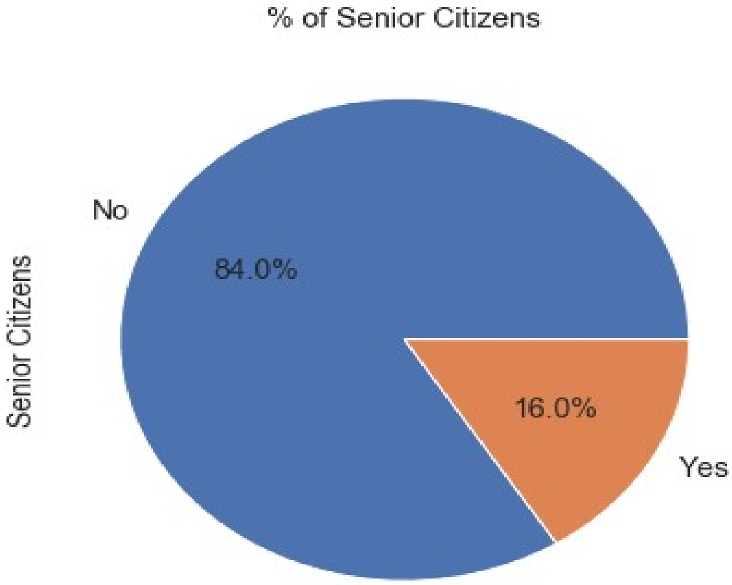

Figure 6.2: % of Senior Citizen

6.3 Partner and dependent status- About 49% of the clients are having partner, while only 30% of the total clients have dependents.

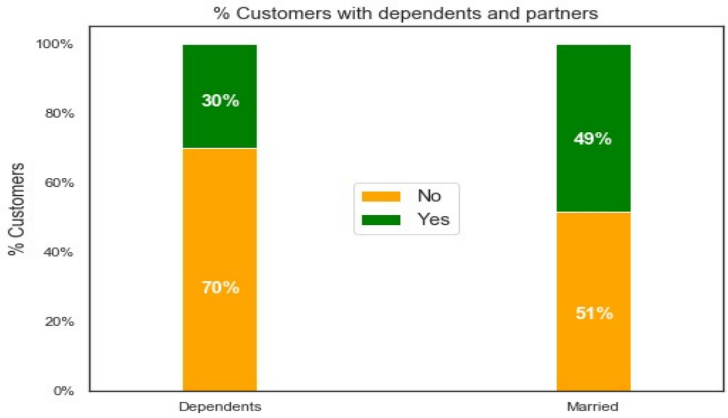

Figure 6.3: % Customers with dependents and partners

6.4 Tenure: After looking at the below histogram we can see that a lot of clients have been with the telecom company for just a month, while quite a many are there for about 72 months. This could be potentially because different clients have different contracts. Thus based on the contract they are into it could be more/less easier for the clients to stay/leave the telecom company.

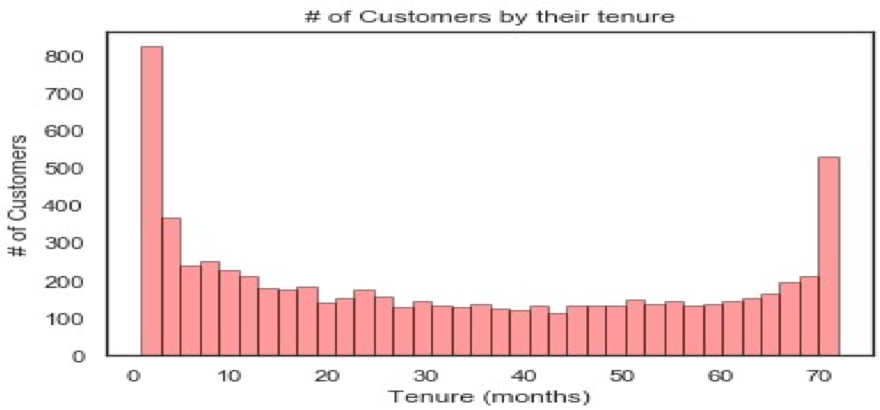

Figure 6.4: % Customers by their tenure

6.5 Contracts: To understand the above graph, let's first look at the % of clients by different contracts. As we can see from this graph most of the clients are in the month to month contract. While there are equal number of clients in the 1 year and 2 year contracts.

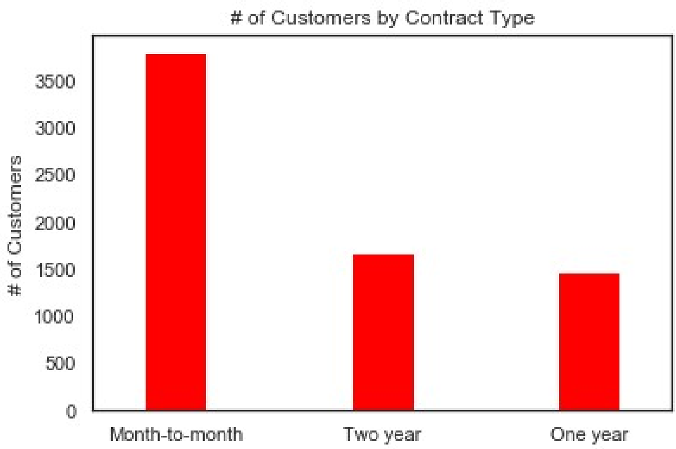

Figure 6.5: % Customers by contract type

6.6 Below we will understand the tenure of clients based on their contract type. Interestingly most of the monthly contracts last for 1-2 months, while the 2 year contracts tend to last for about 70 months. This shows that the clients taking a longer contract are more loyal to the company and tend to stay with it for a longer period of time.

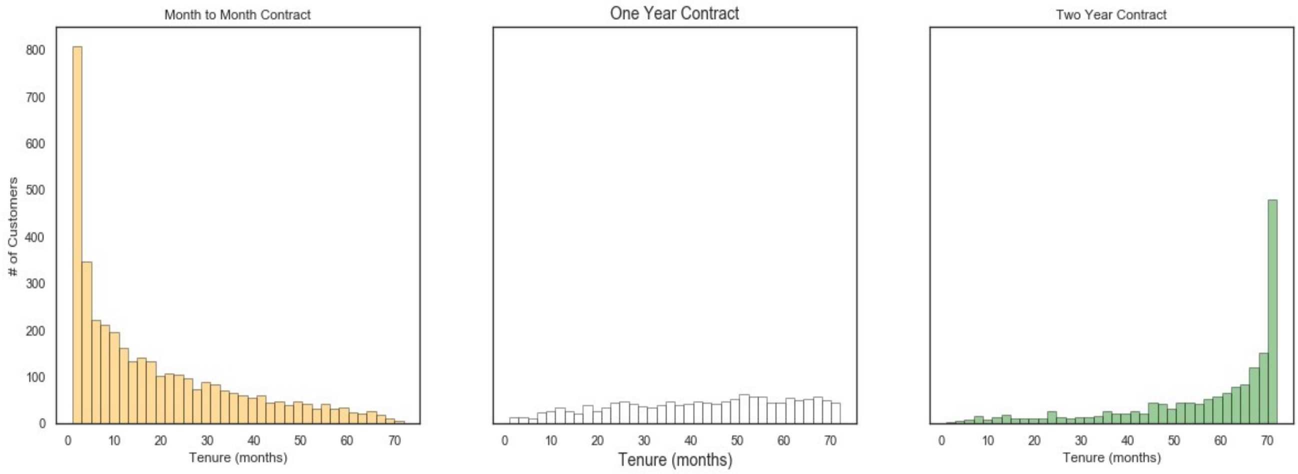

Figure 6.6: Tenure of client based on their contract type

6.7 Let us now look at the distribution of various services used by clients.

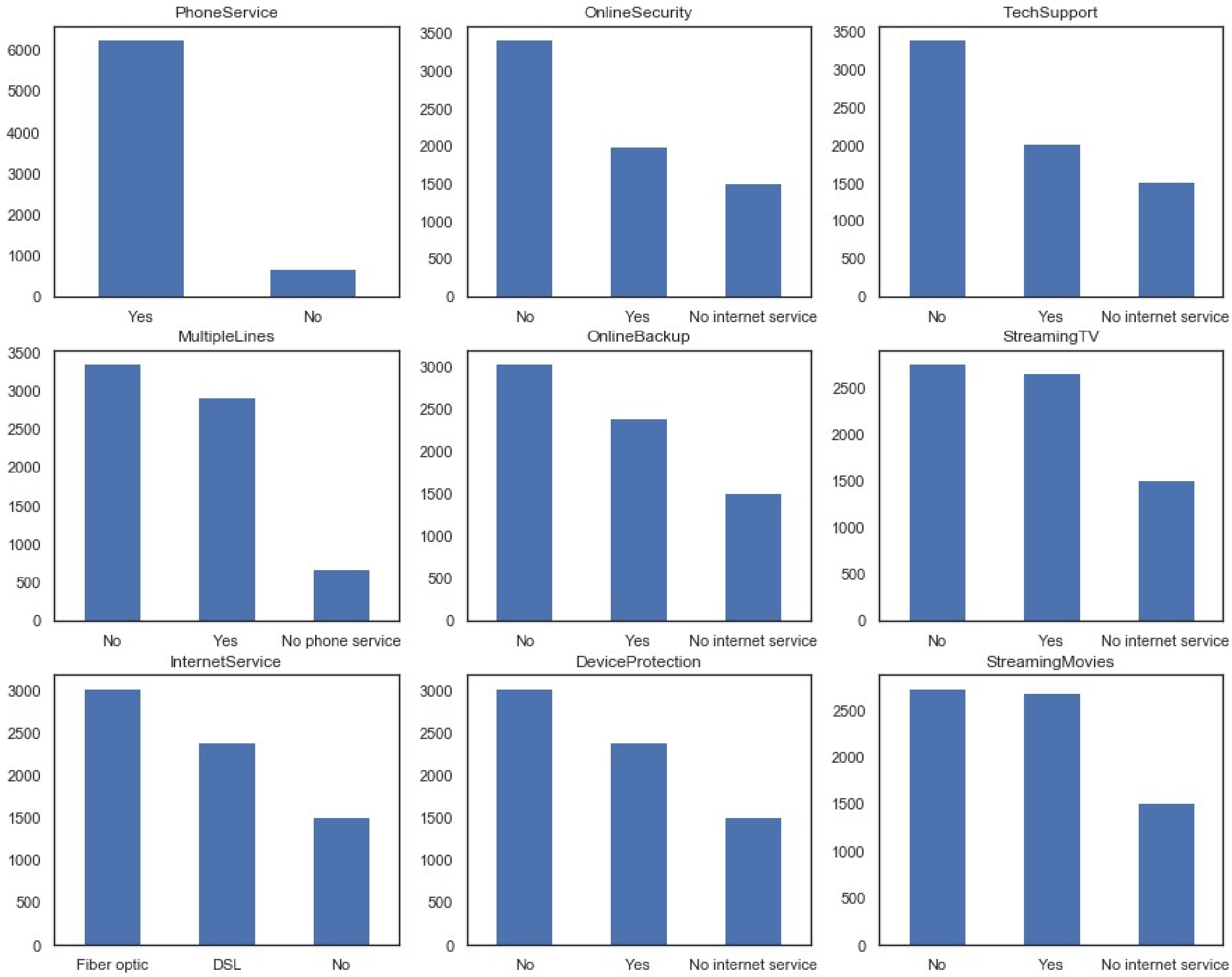

Figure 6.7: Distribution of various services used by clients

6.8 We will observe that the total charges increases as the monthly bill for a client increases.

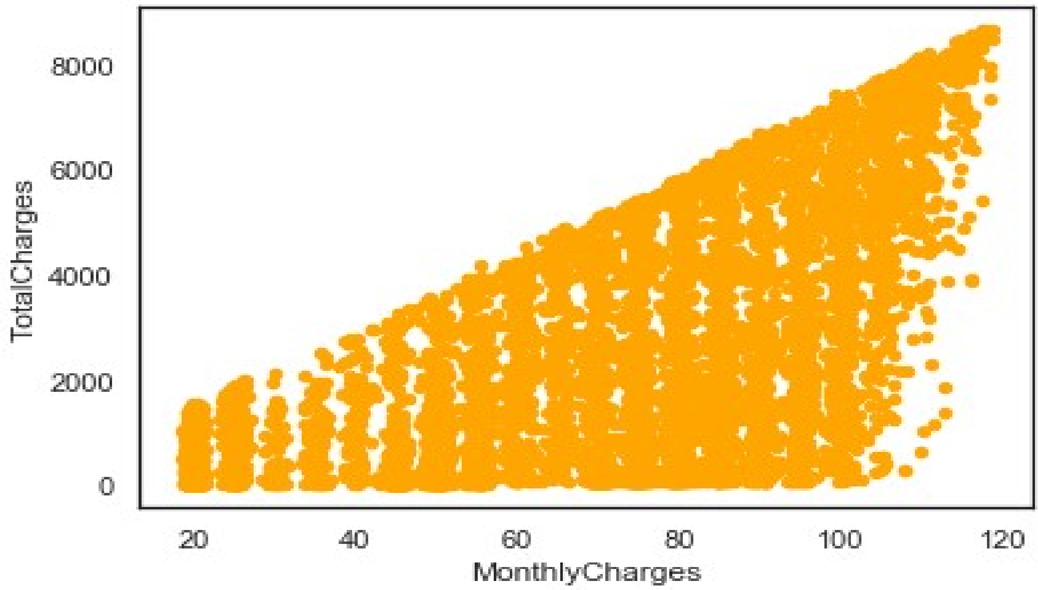

Figure 6.8: Distribution of TotalCharges vs. MonthlyCharges

6.9 Churn Rate: Finally, let's take a look at our predictor variable (Churn) and understand its interaction with other important variables as was found out in the correlation plot.
Let’s first look at the churn rate in our data.

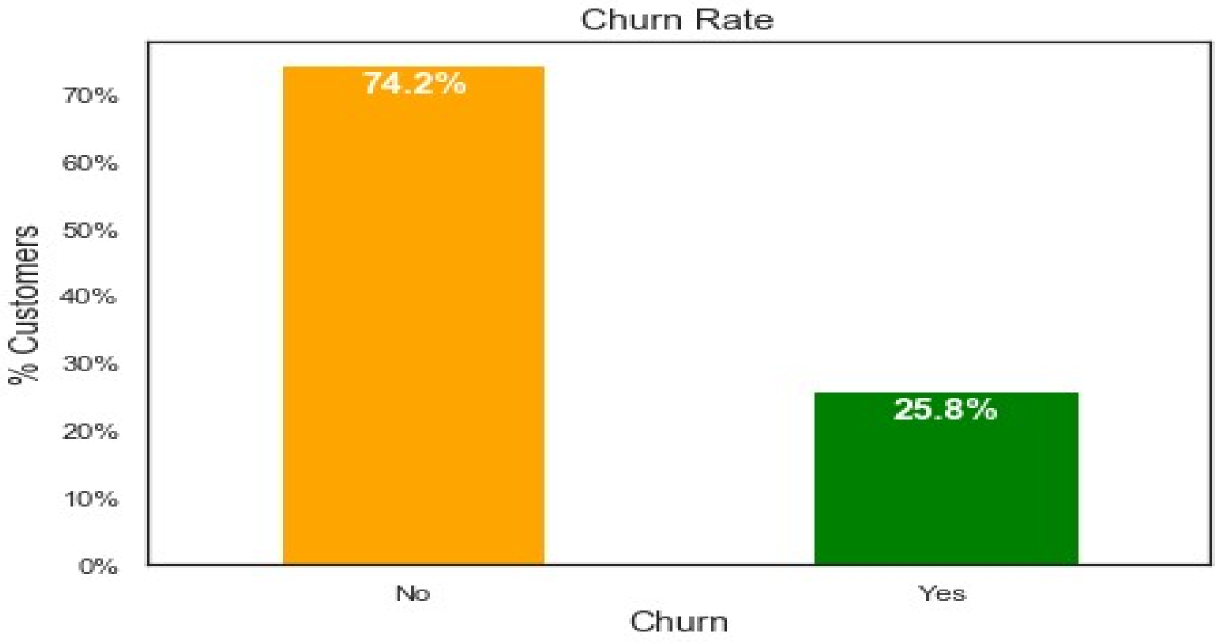

Figure 6.9: Churn Rate

In our data, 74.2% of the clients do not churn. Clearly the data is skewed as we would expect a large majority of the clients to not churn. This is important to keep in mind for our modeling as skeweness could lead to a lot of false negatives. We will see in the modeling section on how to avoid skewness in the data.

6.10 Logistic Regression- It is important to scale the variables in logistic regression so that all of them are within a range of 0 to 1. This helped me improve the accuracy from 79.7% to 80.7%. Further, you will notice below that the importance of variables is also aligned with what we are seeing in Random Forest algorithm and the EDA we conducted above.

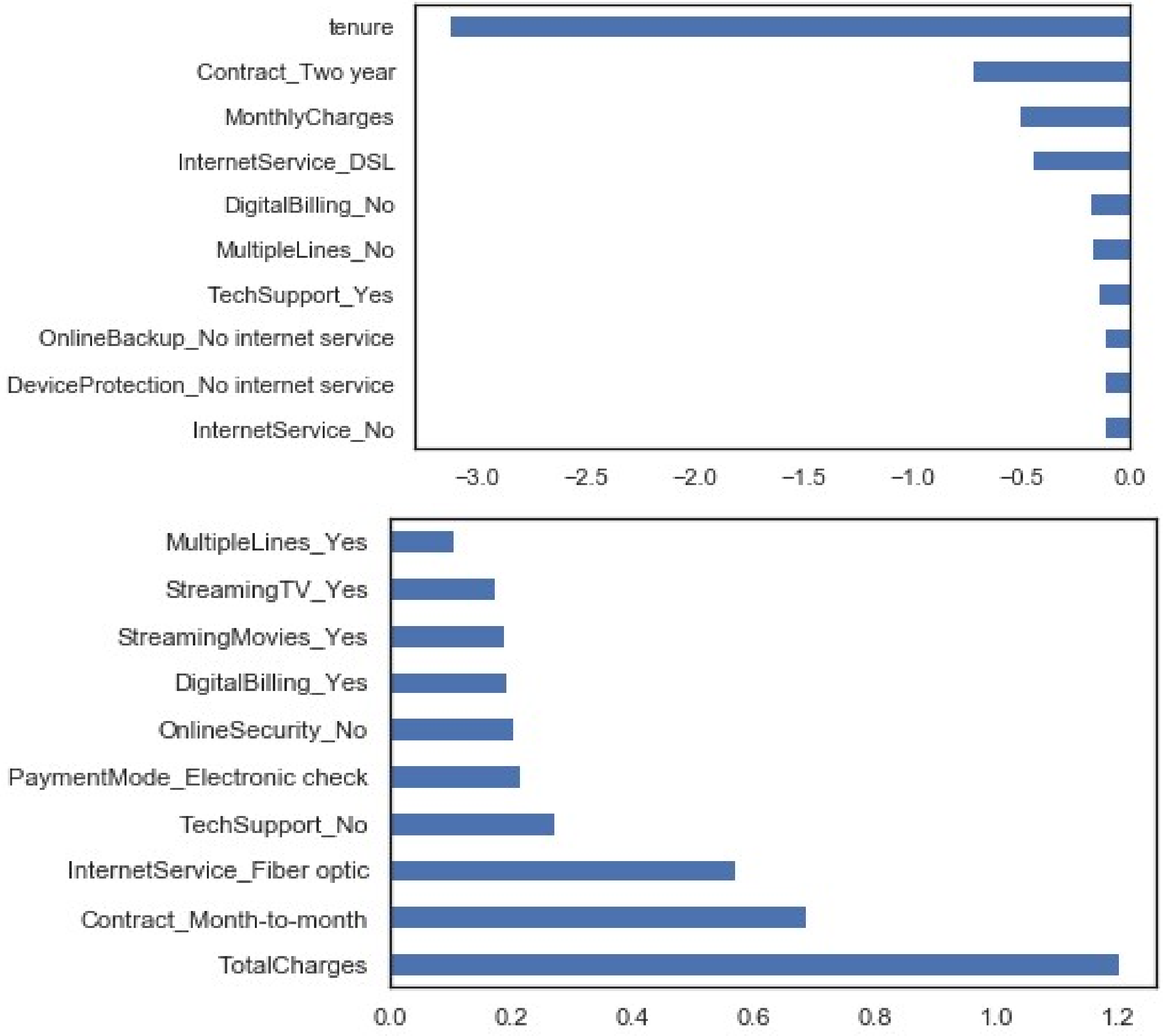

Figure 6.10: weights of all the variables

Observations

We can see that some variables have a negative relation to our predicted variable (Churn), while some have positive relation. Negative relation means that likeliness of churn decreases with that variable. Let us summarize some of the interesting features below:

- As we saw in our EDA, having a 2 month contract reduces chances of churn. 2 month contract along with tenure have the most negative relation with Churn as predicted by logistic regressions
- Having DSL internet service also reduces the probability of Churn
- Lastly, total charges, monthly contracts, fiber optic internet services and seniority can lead to higher churn rates. This is interesting because although fiber optic services are faster, clients are likely to churn because of it.

6.11 Random Forest-

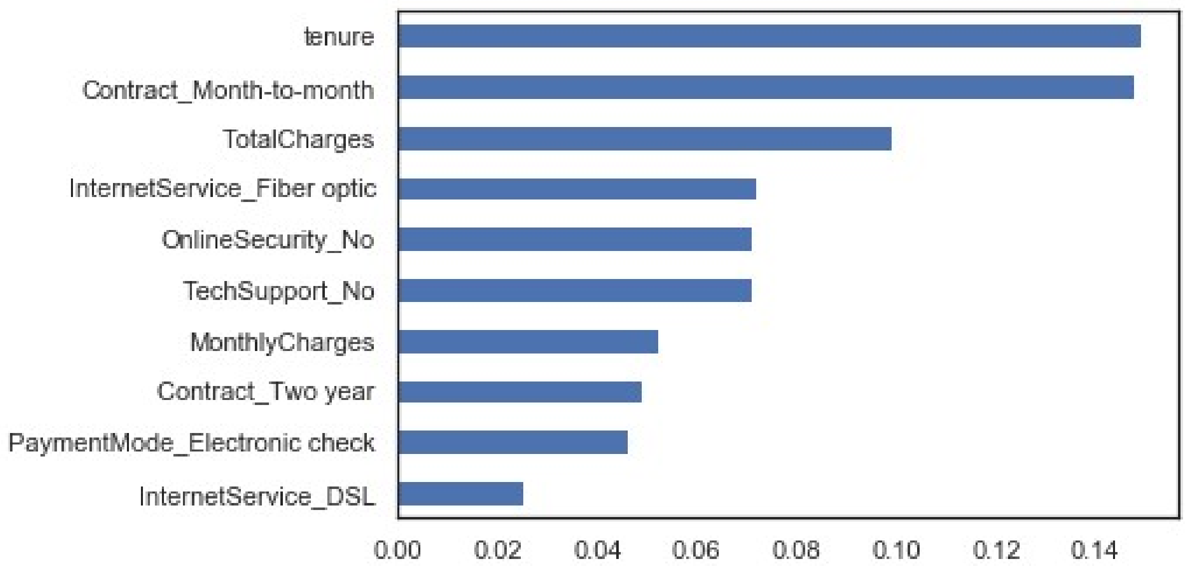

Figure 6.11: weights of all the variables

Observations:

- From random forest algorithm, monthly contract, tenure and total charges are the most important predictor variables to predict churn.
- The results from random forest are very similar to that of the logistic regression and in line to what we had expected from our EDA.

6.12 Support Vector Machine (SVM) - With SVM we were able to increase the accuracy to up to 82%. However, we need to take a deeper look at the true positive and true negative rates, including the area under the Curve (AUC) for a better prediction.

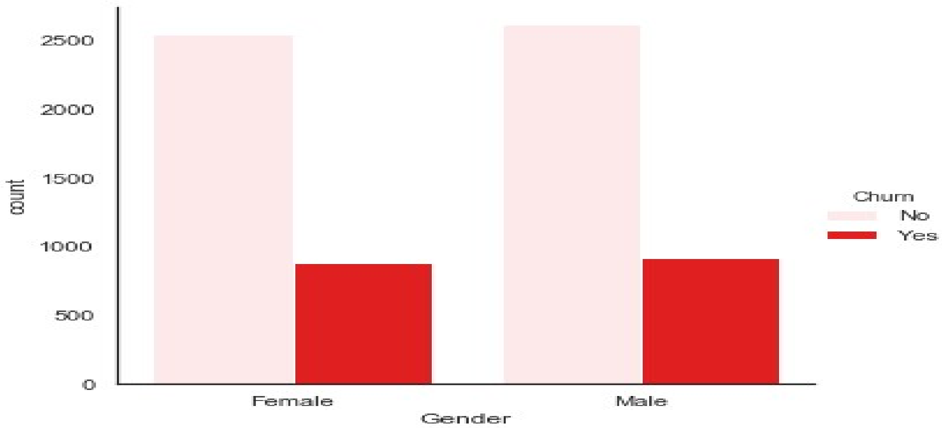

Figure 6.12: Churn Rate with Gender

6.13 XGBoost- Interestingly with XG Boost we were able to increase the accuracy on test data to almost 82%. Clearly, XGBoost is a winner among all other technique s. XG Boost is a slow learning model and is based on the concept of Boosting

Chapter 9

Conclusion and Future Scope

9.1 Conclusion

- Tenure: After observing the histogram Fig-6.4, we can see that a lot of clients have been with the telecom company for just a month, while quite a many are there for about 72 months. This could be potentially because different clients have different contracts. Thus based on the contract they are into it could be more/less easier for the clients to stay/leave the telecom company.
- Contracts: As we can observe from the Fig-6.5, that most of the clients are in the month to month contract. While there are equal number of clients in the 1 year and 2 year contracts. Interestingly most of the monthly contracts last for 1-2 months, while the 2 year contracts tend to last for about 70 months. This shows that the clients taking a longer contract are more loyal to the company and tend to stay with it for a longer period of time.
- The total charges increases as the monthly bill for a client increases.
- Churn vs. Tenure: The clients who do not churn, they tend to stay for a longer tenure with the telecom company.
- Churn by Contract Type: The clients who have a month to month contract have a very high churn rate.
- Churn by Seniority: Senior Citizens have almost double the churn rate than younger age population.
- Churn by Monthly Charges: Higher % of clients churn when the monthly charges are high.
- Churn by Total Charges: It seems that there is higher churn when the total charges are lower.
- Engage and educate the clients.
- Offering incentives to clients likely to churn as per the client interest.
- Providing better service to clients.
- Offering more long term contracts.
- Optimizing pricing strategy.
- Improving campaign management.
- Enhancing client satisfaction.
- Getting timely insights and no delay response.

9.2 Future Scope

The future scope of this project could include conducting extensive experiments for individual category of client attributes, such as client demographic and call data records, rather building a model using all the available predictor variables.

References

[1] Jae-Hyeon Ahn, Sang-Pil Han, Yung-Seop Lee, "Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry", Telecommunications policy, vol. 30, no. 10, pp. 552-568, 2006.

[2] A. Neslin Scott, Sunil Gupta, Wagner Kamakura, Junxiang Lu, Charlotte H. Mason, "Defection detection: Measuring and understanding the predictive accuracy of customer churn models", Journal of marketing research, vol. 43, no. 2, pp. 204-211, 2006.

[3] John Hadden, Ashutosh Tiwari, Rajkumar Roy, Dymitr Ruta, "Computer assisted customer churn management: State-of-the-art and future trends", Computers & Operations Research, vol. 34, no. 10, pp. 2902-2917, 2007.

[4] Hee-Su Kim, Choong-Han Yoon, "Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market", Telecommunications policy, vol. 28, no. 9, pp. 751-765, 2004.

[5] Daihwan Min, Lili Wan, "Switching factors of mobile customers in Korea", Journal of Service Science, vol. 1, no. 1, pp. 105-120, 2009.

[6] V. Umayaparvathi and K. Iyakutti, "Attribute selection and Customer Churn Prediction in telecom industry," 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE), Ernakulam, 2016, pp. 84-90.

STUDENT'S CONTRIBUTION TO THE PROJECT

PROJECT TITLE: CLIENT ATTRITION FORECASTING

NAME OF STUDENT: ADITYA KUMAR LAL
ROLL NO: 1605170

NAME OF STUDENT: AMAN PATNAIK
ROLL NO: 1605174

NAME OF STUDENT: ASHUTOSH SHARMA
ROLL NO: 1605184

NAME OF STUDENT: SUMIT KUMAR SONI
ROLL NO: 1605241

CONTRIBUTION

S No:	**Responsibilities**	**ADITYA**	**AMAN**	**ASHUTOSH**	**SUMIT**
1	Time and Objective Tracking	✓	✓	✓	✓
2	Project Planning	✓	✓	✓	✓
3	Conducting Research	✓	✓	✓	✓
4	Report Writing	✓	✓	✓	✓
5	Preparing Slides	✓	✓	✓	✓
6	Code Implementation	✓	✓	✓	✓
7	Closing	✓	✓	✓	✓

Full Signature of Supervisor: ……………………………

Full signature of the student: ……………………………..