

**A Project Report on
Predicting Telecom Churn using IBM Watson Studio**

**BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE AND ENGINEERING**

**BY
PININTI HARISH REDDY,
C.SAI BHANU PRAKASH,
G.PRAVEEN KUMAR,
BOBBA JAGADISH ,**

**Under the Guidance of
Ms.Pradeepthi**

CERTIFICATE

THIS IS TO CERTIFY THAT **B.JAGADISH**, ENROLLED IN THE B-TECH DEGREE PROGRAM (COMPUTER SCIENCE ENGINEERING) OF THE **SREE DATTHA GROUP OF INSTITUTIONS,IBRAHIMPATNAM.**

C.SAI BHANU PRAKASH, ENROLLED IN THE B-TECH DEGREE PROGRAM (COMPUTER SCIENCE ENGINEERING) OF THE **TKR COLLAGE OF ENGINEERING AND TECHNOLOGY**,**G.PRAVEEN KUMAR** ENROLLED IN THE B-TECH DEGREE PROGRAM (COMPUTER SCIENCE ENGINEERING) OF THE **TKR COLLAGE OF ENGINEERING AND TECHNOLOGY**,

PININTI HARISH REDDY ENROLLED IN THE B-TECH DEGREE PROGRAM (COMPUTER SCIENCE ENGINEERING) OF THE **TKR COLLAGE OF ENGINEERING AND TECHNOLOGY** HAS SUCCESSFULLY COMPLETED THE FOUR WEEK OF INTERNSHIP CUM HANDS-ON TRAINING PROGRAM CONDUCTED BY **SMART BRIDGE** AT STANLEY COLLAGE OF ENGINEERING FOR WOMEN IN 'INTRODUCTION TO MACHINE LEARNING WITH IBM WATSON STUDIO' DURING THE PERIOD OF 3 JUNE,2019 TO 28 JUNE,2019 UNDER THE GUIDANCE OF MS.PRADEEPTHI. DURING THE PERIOD OF INTERNSHIP WITH US HE WAS FOUND PUNCTUAL AND HARDWORKING AND INQUISITIVE .

MENTOR

1 . TELECOM CHURN

1.1 INTRODUCTION

The churn rate is a particularly useful measurement in the telecommunications industry. This includes cable or satellite television providers, Internet providers, and telephone service providers (landline and wireless service providers). As most customers have multiple options from which to choose, the churn rate helps a company determine how it is measuring up to its competitors. If one out of every 20 subscribers to a high-speed Internet service terminated their subscriptions within a year, the annual churn rate for that internet provider would be 5%.

The telecommunications sector has become one of the main industries in developed countries. The technical progress and the increasing number of operators raised the level of competition . Companies are working hard to survive in this competitive market depending on multiple strategies. Three main strategies have been proposed to generate more revenues acquire new customers,

upsell the existing customers, and increase the retention period of customers. However, comparing these strategies taking the value of return on investment of each into account has shown that the third strategy is the most profitable strategy, proves that retaining an existing customer costs much lower than acquiring a new one , in addition to being considered much easier than the upselling To the third strategy, To the third strategy, companies have to decrease the potential of customer's churn, known as "the customer movement from one provider to another".

Customers churn is a considerable concern in service sectors with high competitive services. On the other hand, predicting the customers who are likely to leave the company will represent potentially large additional revenue source if it is done in the early Phase.

Many research confirmed that machine learning technology is highly efficient to predict this situation. This technique is applied through learning from previous data .

We focused on evaluating and analyzing the performance of a set of tree-based machine learning methods and algorithms for predicting churn in telecommunications

companies. We have experimented a number of algorithms such as Decision Tree, Random Forest, to build the predictive model of customer Churn after developing our data preparation, feature engineering, and feature selection Methods.

The data used in this research contains all customers information throughout nine months before baseline. formats which are structured, semi-structured, and unstructured. The data also comes very fast and needs a suitable big data platform to handle it. The dataset is aggregated to extract features for each customer.

We built the social network of all the customers and calculated features like degree centrality measures, similarity values, and customer's network connectivity for each customer. SNA features made good enhancement in AUC results and that is due to the contribution of these features in giving more different information about the Customers.

We focused on evaluating and analyzing the performance of

a set of tree-based machine learning methods and algorithms for predicting churn in telecommunications companies. We have experimented a number of algorithms such as Decision Tree, Random Forest, to build the predictive model of customer Churn after developing our data preparation .

1.2 OBJECTIVES OF RESEARCH

Presented an advanced methodology of data mining to predict churn for prepaid customers using dataset for call details of 7044 customers with 12 features, and a dependent churn parameter with two values: Yes/No. Three machine learning algorithms were used: Support Vector Machine, and Bayes Networks to predict churn factor. The author used AUC to measure the performance of the algorithms. The AUC values were 99.10%, 99.55% And 99.70% for support vector machine

The dataset used in this study is small and no missing values existed.

Firms in telecommunication sector have detailed call records. These firms can segment their customers by using

call records for developing price and promotion strategies. By using machine learning techniques, the subscribers who are intended not to make any payments, can be detected early. And also, financial losses can be prevented. For this type of analysis.

According to usage patterns subscribers are divided into specific clusters. Showing inconsistent features are determined and will be reviewed.

By using Machine learning Techniques, International Roaming Agreements also can be optimized.

Churn prediction and management have become of great concern to the mobile operators. Mobile operators wish to retain their subscribers and satisfy their needs. Hence, they need to predict the possible churners and then utilize the limited resources to retain those customers. Through an analysis result from a telecom provider, the results indicated that the proposed approach has pretty good prediction accuracy by using customer demography, billing information, call detail records, and service changed log to

build churn prediction mode by using machine learning models.

In a large software system knowing which files are most likely to be fault-prone is valuable information for project managers. They can use such information in prioritizing software testing and allocating resources accordingly. To predict defect proneness of modules they trained models on publicly available Nasa MDP data. In their experiments they used Static Call Graph Based Ranking as well as Nearest Neighbor Sampling for constructing method level defect predictors.

1.3 PROBLEM STATEMENT

The study of predicting which persons are going to churn in advance will help the telecommunication industry and the CRM department to identify which persons are going to leave the network. The problem of our work discussed is the classification problem i.e. to classify each subscriber as potential churner or potential non churner. The

framework discussed below is based on the Knowledge Discovery Data (KDD) process. Our framework consists of the following five modules:

❖ Data Acquisition:

Acquiring data from the teleaset industry is a big task because of the fear of misusing it. The data set for this study acquired from the KDD Cup 2009. It is used to analyze the marketing tendency of customers from the large databases from the French Telecom company Orange.

❖ Data Preparation:

Since the dataset acquired cannot be applied directly to the churn prediction models, so aggregation of data is required where new variables are added to the existing variables by viewing the periodic usage behavior of the customers. These variables are very important in predicting the behavior of customers in advance as they contain critical information used by the prediction models.

❖ Data Preprocessing:

Data preprocessing is the most important phase in prediction models as the data consists of ambiguities, errors, redundancy which needs to be cleaned beforehand. The data gathered from multiple sources first is aggregated then cleaned as the complete data collected is not suitable for modeling purposes. The records with unique values do not have any significance as they do not contribute much in predictive modeling. Fields with too many null values also need to be Discarded.

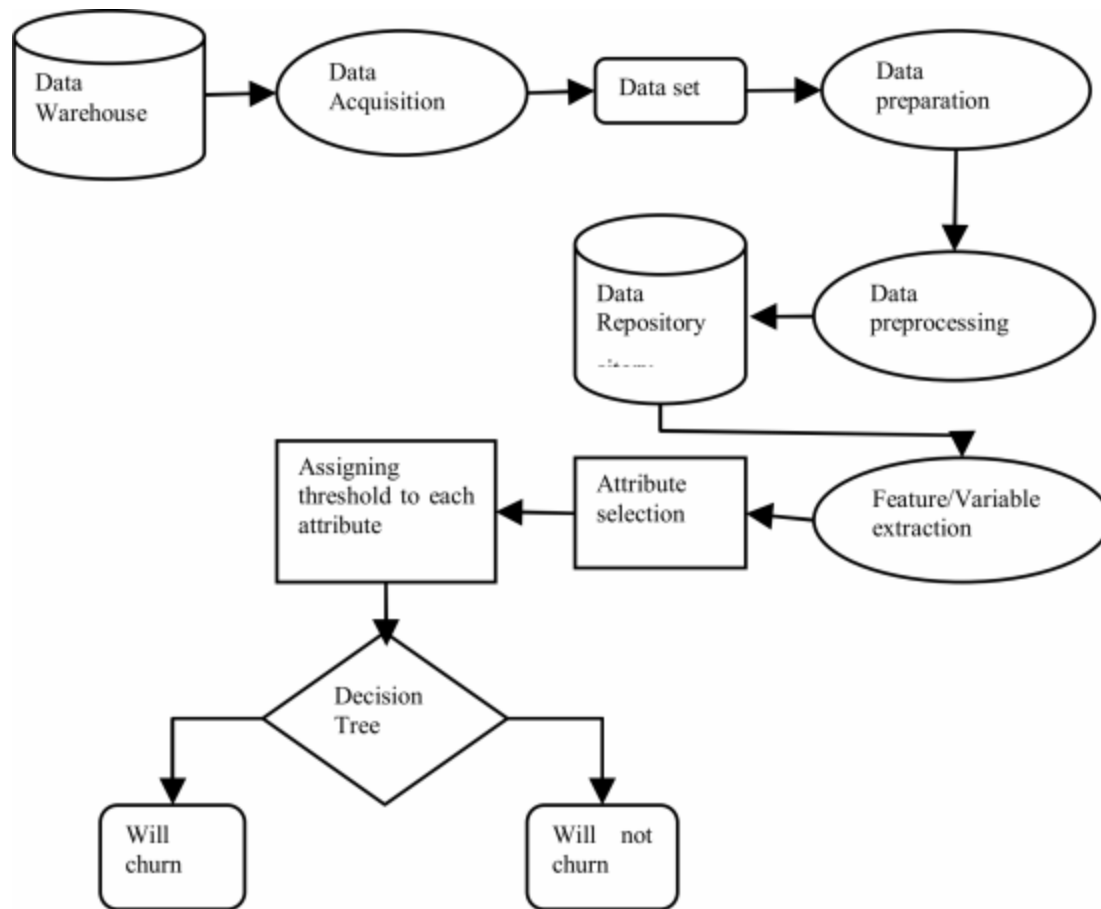
❖ Data Extraction:

The attributes are identified for classifying process. In our work, we have worked with and categorical values.

❖ Decision:

The rule set will let the subscribers identify and classify in the different categories of churners and non churners by setting a particular threshold value.

The framework design used in our work is given in [Figure](#).



1.4 INDUSTRY PROFILE

Telecommunications industry:

Introduction

India is currently the world's second-largest telecommunications market with a subscriber base of 1.20 billion and has registered strong growth in the past decade and half. The Indian mobile economy is growing rapidly and will contribute substantially to India's Gross Domestic Product (GDP), according to report prepared by GSM

Association (GSMA) in collaboration with the Boston Consulting Group (BCG). As of January 2019, India has witnessed a 165 per cent growth in app downloads in the past two years.

The liberal and reformist policies of the Government of India have been instrumental along with strong consumer demand in the rapid growth in the Indian telecom sector. The government has enabled easy market access to telecom equipment and a fair and proactive regulatory framework that has ensured availability of telecom services to consumer at affordable prices. The deregulation of Foreign Direct Investment (FDI) norms has made the sector one of the fastest growing and a top five employment opportunity generator in the country.

Market Size

With 560.01 million internet subscribers, as of September 2018, India ranks as the world's second largest market in terms of total internet users.

Further, India is also the world's second largest telecommunications market with 1,197.87 million subscribers, as of December 2018.

Moreover, in 2017, India surpassed USA to become the second largest market in terms of number of app downloads.

The country remained as the world's fastest growing market for Google Play downloads in the second and third quarter of 2018.

Over the next five years, rise in mobile-phone penetration and decline in data costs will add 500 million new internet

users in India, creating opportunities for new businesses.

Investment/Major development

With daily increasing subscriber base, there have been a lot of investments and developments in the sector.

The industry has attracted FDI worth US\$ 32.45 billion during the period April 2000 to December 2018, according to the data released by Department of Industrial Policy and Promotion (DIPP).

Some of the developments in the recent past are:

- ❖ During the first quarter of 2018, India became the world's fastest-growing market for mobile applications. The country remained as the world's fastest growing market for Google Play downloads in the second and third quarter of 2018.**
- ❖ Bharti Airtel is planning to launch 6,000 new sites and 2,000 km of optical fiber in Gujarat in 2018-19.**
- ❖ The number of mobile wallet transaction increased 5 per cent month-on-month to 325.28 million in July 2018.**
- ❖ As of June 2018, BSNL is expected to launch its 5G services by 2020.**
- ❖ Vodafone India and Idea Cellular have merged into 'Vodafone Idea' to become India's largest telecom company, as of September 2018.**

Government Initiatives

The government has fast-tracked reforms in the telecom sector and continues to be proactive in providing room for growth for telecom

companies. Some of the other major initiatives taken by the government are as follows:

- ❖ **The Government of India is soon going to come out with a new National Telecom Policy 2018 in lieu of rapid technological advancement in the sector over the past few years. The policy has envisaged attracting investments worth US\$ 100 billion in the sector by 2022.**
- ❖ **The Department of Information Technology intends to set up over 1 million internet-enabled common service centres across India as per the National e-Governance Plan.**
- ❖ **FDI cap in the telecom sector has been increased to 100 per cent from 74 per cent; out of 100 per cent, 49 per cent will be done through automatic route and the rest will be done through the FIPB approval route.**
- ❖ **FDI of up to 100 per cent is permitted for infrastructure providers offering dark fibre, electronic mail and voice mail.**
- ❖ **The Government of India has introduced Digital India programme under which all the sectors such as healthcare, retail, etc. will be connected through internet**

Achievements

Following are the achievements of the government in the past four years:

- ❖ **Department of Telecommunication launched 'Tarang Sanchar' - a web portal sharing information on mobile towers and EMF Emission Compliances.**
- ❖ **Six-fold increase in Government spending on telecommunications infrastructure and services in the**

country – from Rs 9,900 crores (US\$ 1.41 billion) during 2009-14 to Rs 60,000 crores (US\$ 8.55 billion) (actual + planned) during 2014-19.

❖ Over 75 per cent increase in internet coverage – from 251 million users to 446 million

❖ Country-wide Optical Fibre Cable (OFC) coverage doubled – from 700,000 km to 1.4 million km

❖ Five-fold jump in FDI inflows in the Telecom Sector – from US\$ 1.3 Billion in 2015-16 to US\$ 6.1 billion in 2017-18 (up to December 2017)

2.Review of Literature

❖ Companies in telecommunication sector have detailed call records information. These firms can segment their customers by using call records for developing price and promotion strategies.

❖ By using machine learning models techniques , data analytics the users who are intended not to make any payments, can be detected easily and before. And also, financial losses can be prevented. For this type of analysis, Different models are is applied. According to usage patterns subscribers are divided into specific clusters.

Showing inconsistent features are determined and will be reviewed.

**❖ By using models with different Techniques,
International Roaming Agreements also can be
optimized.**

**❖ Machine learning algorithms and knowledge discovery
framework have been successfully applied in a number of
application domains including commerce, astronomy,
geological survey, security, and telecommunications.**

**❖ There are some clustering methods, based on genetic
algorithm for telecommunication customer subdivision.
First, the features of telecommunication customers (such
as the calling behaviour and consuming behaviour) are
extracted. Then, the similarities between the
multidimensional feature vectors of telecommunication
customers are computed and mapped as the distance
between samples on a two-dimensional plane. Finally, the
distances are adjusted to approximate the similarities
gradually by genetic algorithm.**

**❖ Churn prediction and management have become of
great concern to the mobile operators. Mobile
operators wish to retain their subscribers and satisfy**

their needs. Hence, they need to predict the possible churners and then utilize the limited resources to retain those customers. In response to the difficulty of churner prediction, Changes study applies data mining to build a model for churner prediction.

❖ This model helps the industries to take the business decisions smart based on the outcomes from the models and according to which they can prepare strategies according.

❖ Current research also fails to acknowledge the expensive problem of misclassifying non-churners as churners. In addition, most research efforts base their analysis on customer demographic and usage data that can breach governing regulations. It is proposed in this research that customer complaints and repairs data could prove a suitable alternative. We have proposed to build a model for churn prediction for telecommunication companies using data mining and machine learning techniques namely logistic regression and decision trees.

❖ A comparison is made based on efficiency of these

algorithms on the available dataset this model helps a lot in giving more number of offers to the customers who are using there network consistently and honestly so that they can they increase their revenue.

❖ During the estimation of possible churns, data from the previous churns might be used. An efficient churn predictive model benefits companies in many ways. Early identification of customers likely to leave may help to build cost effective ways in marketing strategies. Customer retention campaigns might be limited to selected customers but it should cover most of the customer.

❖ Incorrect predictions could result in a company losing profits because of the discounts offered to continuous subscribers. Therefore, the right predictions of the churn customers has become highly important for the companies, The prominent role that telecommunication sector has come to occupy worldwide makes it all the more important to develop prediction mechanisms along the lines of churn prediction. Few statistics show the importance of the customer retains in this sector

3. DATA COLLECTION

❖ **we have collected this data set in various open source platform where the people can download the dataset the analyze the data and submit the report of various hidden information according that ,they will play the business tactics according the market competition . The research demonstrates that customers can be placed into one of several profiles clusters according to their interactions with the service provider. Based on this, an estimate is possible regarding when the customer can be expected to terminate his/her service with the company.**

❖ **Churn data collection, also known as customer attrition, occurs when customers stop doing business with a company. The companies are interested in identifying segments of these customers because the price for acquiring a new customer is usually higher than retaining the old one.we find one that done the research , For example, if telecom network knew a segment of customers who were at risk of churning they could proactively engage them with special offers**

instead of simply losing them.

❖ **we collected a simple customer churn prediction model using Telco Customer Churn dataset. We chose a random forest to model churned customers, pandas for data crunching and matplotlib for visualizations. We will do all of that above in Python.**

❖ **The code can be used with another dataset with a few minor adjustments to train the baseline model. We also provide few references and give ideas for new features and improvements.**



What data do I need?

Here, one must gather information about past behaviors. One does this by observing a large number of customers and identifying patterns. One needs information on a number of factors that lead us to believe, with confidence, that a certain customer is going to churn out. These factors generally vary by type of business, and include:

1. **Weather less or more subscription period of time?**
2. **Were there regular complaints to customer care that were not satisfactorily resolved?**

4.METHODOLOGY

The dataset consists of 14 variables in all. A few are continuous, rest are categorical. The control variable was customer churn with 2 levels Y/N (i.e.customer has left or not). Initially, we started out with some basic EDA& some visual plots that would help us understand the data better. Since many machine learning algorithms cannot operate on categorical variables, we had to convert this data into numerical variables. We decided to use the Label Encoder method to convert the variable type. After the initial analysis, we proceeded to use techniques like Feature Importance & Feature Selection to see if we had any variables that were redundant & could be discarded in the process of building the models. We decided to use models like svm ,knn. Classifier for this analysis. We split the data into train & test & built a model using each of these classifiers. We used K-fold Cross-validation to evaluate the quality of the predictive models by partitioning the original data

into a training set to train the model, and a test set to evaluate it. We also plotted the ROC curves to check the performance of the binary classifiers.

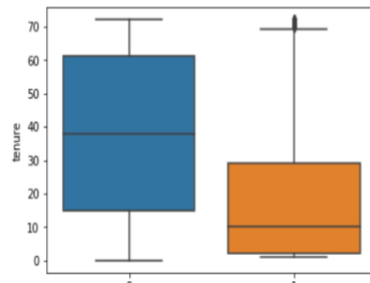
4.1 EXPLORATORY DATA ANALYSIS

We started out with some basic EDA. We have around 7000 observations. from our dataset. We have 12 features & 1 targetvariable. Only 3 features out of 13 were numeric.All the rest were categorical variables. Our customer churn v/s stay data split is in the ratio of 1:3. Ratio of males to females is around 50:50. When we plotted the Churn variable against the customer's tenure, one interesting observation was that most of the customers who leave the telco provider, usually do it within the first year. Beyond the first year, they tend to stick around. Another interesting bit is that their biggest customer base are their oldest customers followed by the newest. Obviously, we would not like to have multicollinearity in our dataset. So, we plotted A heatmap/correlation plot for all the variables. The variables with the highest positive correlation are TotalCharges, MonthlyCharges& Tenure. Looking at

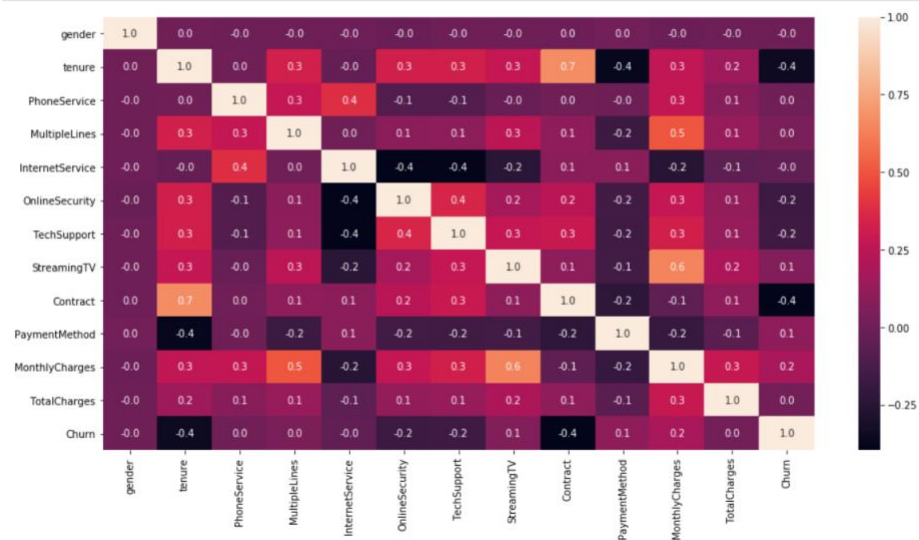
**the data, we figured that TotalCharges is nothing but
Tenure times.**

```
In [13]: sns.boxplot(x='Churn', y='tenure', data=dataset1)
```

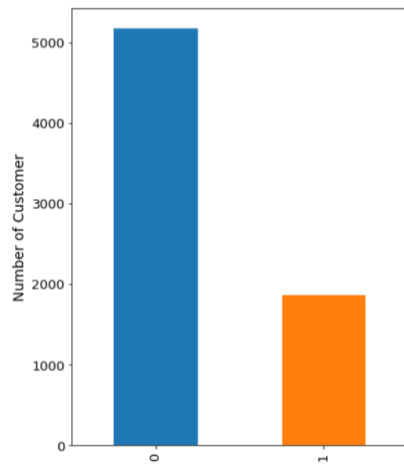
```
Out[13]: <matplotlib.axes._subplots.AxesSubplot at 0x262769259b0>
```



```
In [19]: plt.figure(figsize = (16, 8))  
sns.heatmap(dataset1.corr(), annot=True, fmt=".1f")  
plt.show()
```

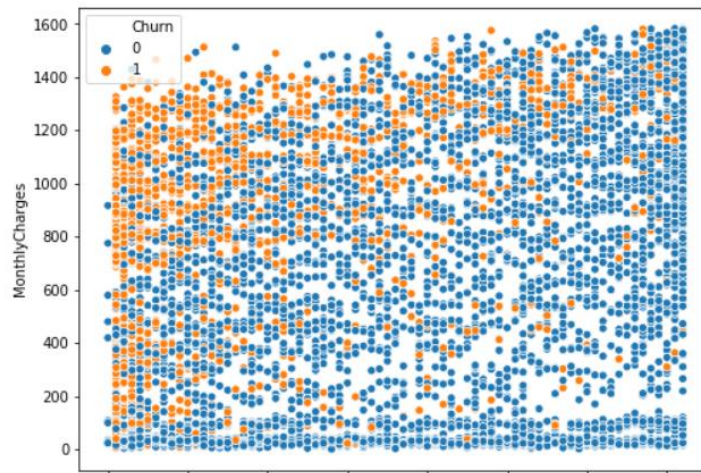


```
In [22]: ax = dataset1["Churn"].value_counts().plot(kind='bar', figsize=(6, 8), fontsize=13)
ax.set_ylabel("Number of Customer", fontsize=14);
```



```
In [24]: plt.figure(figsize=(8, 6))
sns.scatterplot(x = 'tenure', y = 'MonthlyCharges', hue="Churn", data = dataset1)
```

Out[24]: <matplotlib.axes._subplots.AxesSubplot at 0x262774560b8>

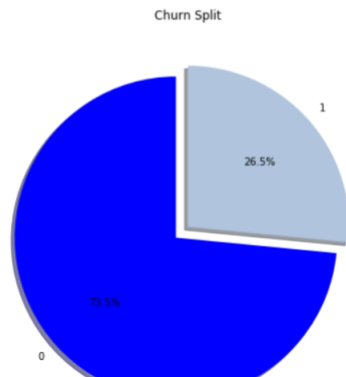



```
In [20]: from collections import Counter

labels, values = zip(*Counter(dataset1["Churn"]).items())
colors = ['blue', 'lightsteelblue']
piechart_df = (pd.DataFrame(list(values),list(labels)))
piechart_df = piechart_df.reset_index()

fig = plt.figure(figsize=[6, 6])

plt.pie(piechart_df[0],labels=piechart_df["index"],startangle=90,explode=(0.1,0),autopct="%1.1f%%", shadow=True, colors=colors)
plt.tight_layout()
plt.title("Churn Split")
plt.show()
```



4.2 MODEL BUILDING

We used quite a few models to check which fits best on our data. Models used are –

- ☐ **Logistic Regression**
- ☐ **Random Forest Classifier**
- k-NN Classifier**

Our target variable is a binary variable. The customer either stays or leaves. Hence, we decided to use the Logistic Regression as our first model to check how the model fits the data. To check the model's predictive performance, we've used the K-fold Cross Validation across

almost each and every classifier barring Random Forest (since that is essentially what a Random Forest does). For the kNN classifier, we initially used k=3 to predict our target label. However, since we did not know what the optimal value should be for k we tried to fit the model against all odd k's between 1 & 50. The model with k =27 had the highest accuracy i.e. the lowest optimal value should be for k we tried to fit the model against all odd k's between 1 & 50. The model with k = 27 had the highest accuracy .

4.3 EVALUATION METRICS

CONFUSION MATRIX

```
In [31]: #SVM
from sklearn.svm import SVC
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.5, random_state = 10)

classifier = SVC(kernel = 'rbf')
classifier.fit(x_train, y_train)
y_pred = classifier.predict(x_test)

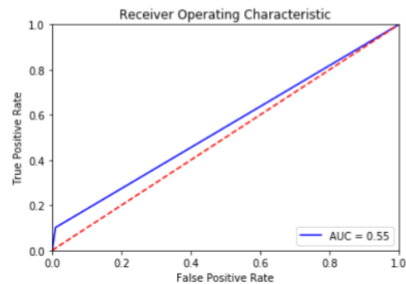
import sklearn.metrics as metrics
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
cm
```

C:\ProgramData\Anaconda3\lib\site-packages\sklearn\svm\base.py:196: FutureWarning: The default value of gamma will change from 'auto' to 'scale' in version 0.22 to account better for unscaled features. Set gamma explicitly to 'auto' or 'scale' to avoid this warning.
"avoid this warning.", FutureWarning)

```
Out[31]: array([[2589,  26],
               [ 815,  92]], dtype=int64)
```

ROC CURVE

```
In [35]: #plotting the roc scalar
plt.title('Receiver Operating Characteristic')
plt.plot(fpr, tpr, 'b', label = 'AUC = %0.2f' % roc_auc)
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1], 'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('True Positive Rate')
plt.xlabel('False Positive Rate')
plt.show()
```

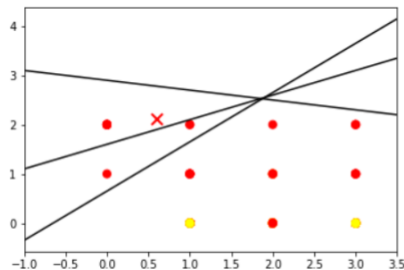


SCATTER PLOT

```
In [37]: xfit = np.linspace(-1, 3.5)
plt.scatter(x[:, 1], x[:, 8], c=y, s=50, cmap='autumn')
plt.plot([0.6], [2.1], 'x', color='red', markeredgewidth=2, markersize=10)

for m, b in [(1, 0.65), (0.5, 1.6), (-0.2, 2.9)]:
    plt.plot(xfit, m * xfit + b, '-k')

plt.xlim(-1, 3.5);
```



5.Findings And Suggestions

Have you been in a situation where you have lost customers to some other company? A large number of organizations today rely heavily on retaining their existing customer base and a large part of the revenue comes in from this sector. All industries suffer with “Churn”. It is seen more in the

Satellite TV, telephone and internet providers where the percentage of people switching from one provider to another is significant month to month. A “churn” with respect to the Telecom industry, is defined as the percentage of subscribers moving from a specific service or a service provider to another in a given period of time. Research shows today that the companies these companies have an average churn of 1.9 to 2 percent month on month and annualized churn ranging from 10 to 60 percent.

Reasons for a Churn

An organization loses its customers to its competition for various reasons. What mainly attracts customers to go ahead with a shift is attributed to price of the product and its quality. This is true for a Telecom industry as well.

Churn of 10 to 60 percent in any company is a big number which can affect the company’s overall growth. The reputation of the company also goes down in the market if the percentage churn increases year on year. Future business gets affected and in turn your entire sales engine fails. Many companies have gone bankrupt and incurred

heavy losses due to a large percentage of churn. You lose not only revenue generated by these customers but also the resources you have invested for them. However, for the customer to be retained, it is very important to also measure customer satisfaction. Many researches show that plenty of customers switch to a different provider because of the lack of satisfaction. Value added service is another reason for Churn. Telecom companies have started a new offering called Triple play combining the TV, broadband and the phone offering as compared to the traditional model of just the phone services. This is seen as a value add to retain customers. The Triple play not only helps retain customers but also increases the Average revenue per user (ARPU) directly contributing to the revenue of the company.

Tips to reduce Churn

- “After sales” service is a key to retain customers.**

Telco’s should categorize their customers based on the ARPU into different buckets and should have the privileged support based on the category as done in the banks today. This will have an improved customer satisfaction as the users will experience a special

privilege.

- **Customers should be communicated with the new services offering based on the usage analysis and trends and should be given proactive information on the plans which will benefit the customer.**

- **Now almost all providers are offering the Triple play service as a combined offering and there is a need for the Telco's to look at an alternative for the value added service. Telco's have to come out with new value added services with differentiators to retain customers. The Triple play effect is no longer going to be a value add and will not work as a differentiation in the market.**

- **Effective communication is one way to reduce churn. Being proactive in addressing difficulties and issues faced by your customers not only helps in building trust and reliability but also ensures a strong working Relationship.**

6.conclusion

The importance of this type of research in the telecom market is to help companies make more profit. It has become known that predicting churn is one of the most important sources of income to telecom companies.

Hence, this research aimed to build a system that predicts the churn of customers in SyriaTel telecom company. These prediction models need to achieve high AUC values. To test and train the model, the sample data is divided into 70% for training and 30% for testing. We chose to perform cross-validation with 10-folds for validation. We have applied feature engineering, effective feature transformation and selection approach to make the features ready for machine learning algorithms. In addition, we encountered another problem: the data was not balanced. Only about 5% of the entries represent customers' churn. This problem was solved by undersampling or using trees algorithms not affected by this problem. Four tree based algorithms were chosen because of their diversity and applicability in this type of prediction. These algorithms are Decision Tree, Random Forest. The method of preparation and selection of features and entering the mobile social network features had the biggest impact on the success of this model, since the value of AUC reached 93.301%.

7. Bibliography and References

1.Gerpott TJ, Rams W, Schindler A. Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market. Telecommun Policy.

<https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0191-6#Bib1>

2. Telecommunication subscribers' churn prediction model using machine learning. In: Eighth international conference on digital information management.

<https://journals.sagepub.com/doi/10.1509/jmr.13.0483>