Wireless companies today measure voluntary churn by a monthly figure, such

INTR ODUCTION

as 1.9 percent or 2.1

**Churn prediction** is one of the most

popular Big Data use cases in business. It consists of detecting customers who are l.ikely to cancel a subscription to a

service.

**Churn** is a problem for telecom companies because it is more expensive to acquire a new customer than to keep your existing one from leaving.

percent.

\_P\_R\_OJ\_E\_CT O\_BJ\_E\_CT\_I\_V\_j== :r

methods, lower prices,

building satisfactory

customer relationships,

);:-To predict Customer Churn.

);:- Highlighting the main variables/factors influencing Customer Churn.

);;:- Use various ML algorithms to build prediction models, evaluate the accuracy and performance of these models.

);;:- Finding out the best model for our business case & providing executive summary.

better marketing and, above all, successful customer communications.

DATASET D E SCRIPTION

);;> Source dataset is in csv format.

);;> Dataset contains 7043 rows and 14 columns

*);;>* There is no missing values for the provided input dataset.

*);;>* Churn is the variable which notifies whether

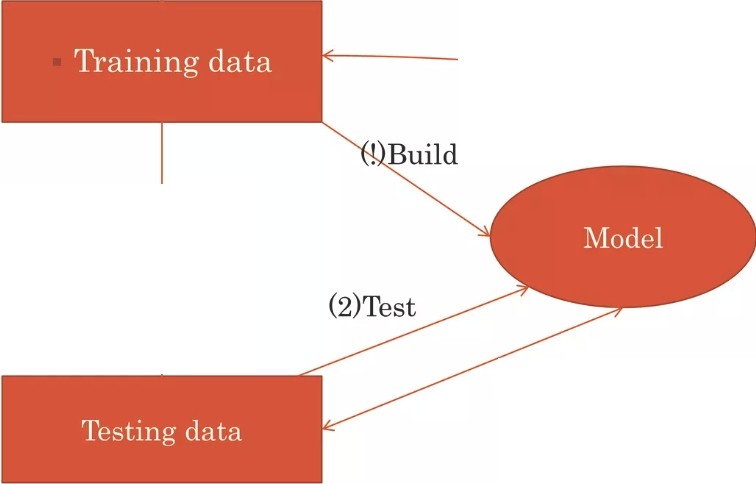
a particular customer is churned or not.

And

e

we will be developing our models to predict

**CHURN PREDICTION MODEL**



(5)Update

inde endent (4)Predict



Real

(3)Test=Ok

**METHODOL OGIES**

EDA(Exploratory Data Analysis): The dataset consists of 12 variables in all. A few are continuous, rest are categorical. The control variable was customer.

Model building which includes defining the purpose if model, determine the model boundary,. build the model, create an interface and export the model.

Evaluating machine learning algorithm is an essential part of project.

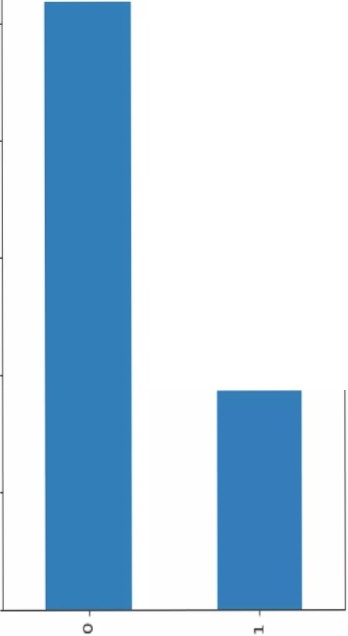
EXPLORATORY DATA ANALYSIS

Data visualisation using seaborn and matplotlib

Exploratory data analysis (EDA) is an approach to analyse data sets & to summarize their main characteristics, often with visual methods.

A Statistical model can be used or not, but primarily EDA is for seeing what the data can tell us beyond the formal modelling or hypothesis.

ax = datasetl["Churn"].value\_counts().plot(kind= 'bar ', figsize=(6, 8), fontsize=13} ax.set\_ylabel("Nunber of Custol'IE!r", fontsize•l4); ax.set\_xlabel("Churn", fontsize•i4);



1:ExitngUser

O: Not E.xitng User

Chum

5000

I..

Ill

E

.

0,

.u...

0

I..

IV

.0

4000

3000

BAR **GRAPH**

Plot Shows that the Users from the Data are likely to be Continuing their

§ 2000

z

1000

0

Subscription

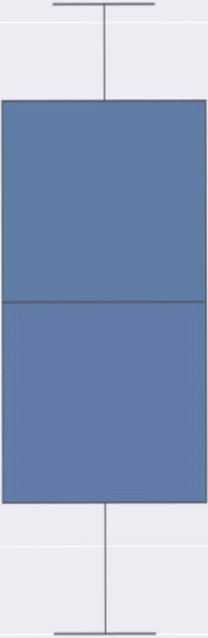
plan(>70%)

•

In (72]: sns.boxplot(x •·curn' , y sns.set(rc { figure.fig

|  |  |
| --- | --- |
| •'tenu,e' | |
| size' : | (10, |

JO



0



0

, data =dataset1) 8)})

Olum

**BOX PLOT**

* *Y.l.e* found outliers in exiting customers which is out of whiskers. An Outlier is an observation that is numerically distant from the rest of the data.
* Using SkewO method we found that Churn data is inconsistent

with tenure

* Customers who disconnecting their subscription plans are selecting short tenure Telecom Company need to offer better plans for those customers who choose short tenures.

•

plt .figure(figsire - (14, 8))

sns.heatmap(datasetl .corr(),annot•True,fllt= .lf")

plt .show()

- 100

**HEAT** MAP

Correlation: Dependence or association is any statistical relationship, whether

causal or not, between two random variables or bivariate data.

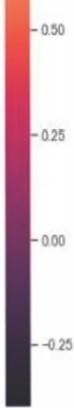
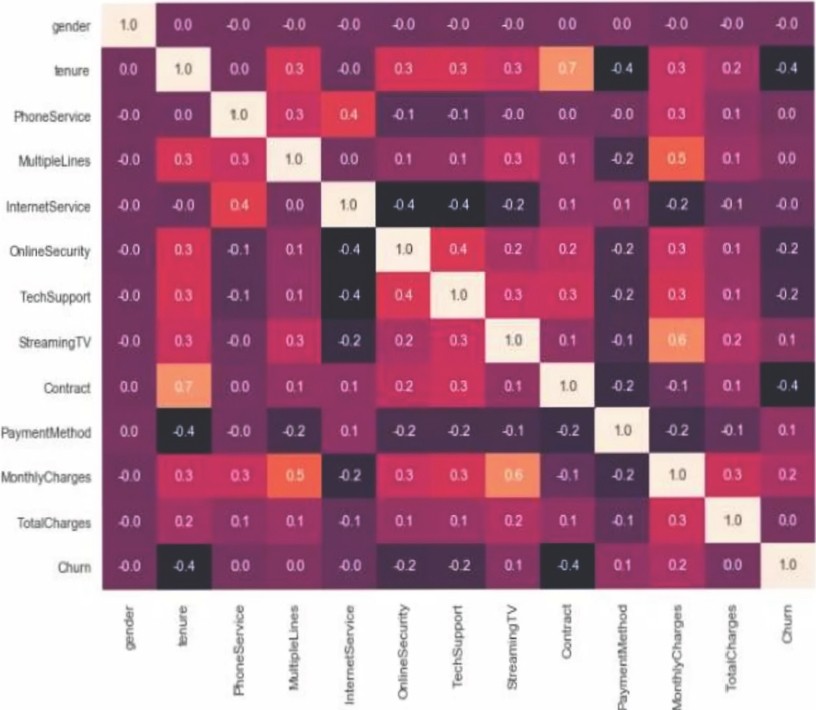
With the help of Correlation matrix, we can find interdependency between variables

l)Least dependency of variables for predicting churn are tenure and contract.

1. Churn variable is

depending more on monthly

charges.



-075

•

labels,values •zip(•Counter (datasetl ["gender"]).ites()) colors ('11ediu1tpurple', 'lightpink ')

piechart\_df •(pd .DataFra111e (list(values),list(labels))) piechart\_df •piechart\_df .reset\_index ()

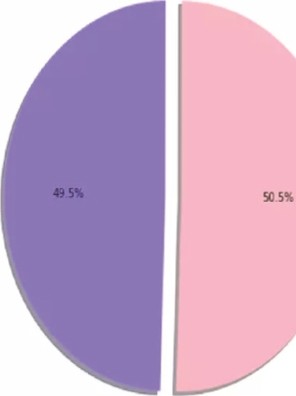
piechart\_df

fig ;plt.figure(figsize•[6, 6))

plt.pie (piechart\_df(0),labels•piechart\_df (''index"],startangle=90,explodes(0.1,0),autopct•"Xl.lf\X ,shadow True, colors=colors: plt .tight\_layout()

plt.title("Gender Split") plt.show()

Gender Split



0

**PIE PLOT**

* 1. **Female subscribers**

**are 49. 5% of the total**

**2)Male subscribers are 55.5% of the total**

•

rom collections import Counter

abels, values = zip(•counter(dataset1["Churn")).items()) olors s ['slateblue ', 'lightsteelblue ']

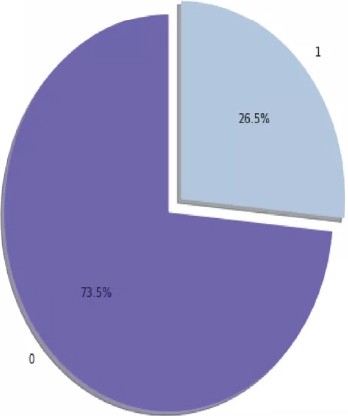
iechart\_df = (pd.DataFrame(list(values),list(labels))) iechart\_df = piechart\_df. reset\_index ()

ig :plt .figure(figsize [6, 6])

lt .pie(piechart\_df[0],labels=piechart\_df["index"],startangle=90,explode=(0.1,0),autopct ="X1.lfXX" , shadol4=True, colors=colors) lt.tight\_layout()

lt.title("Churn Split") lt.show()

Churn Split



PIE PLOT

According to the collected data information 73.5% customers are continuing their subscription

and 26.5% customers disconnected .

•

H plt.figure(figsi e=(8, 6}}

sns.scatterplot(x *=* 'tenure', *y =* 'l lom:hlyCharges',hue="Churr" ,data - dataset )

<matplotlib .axes.\_subplots.AxesSubplot at 0x26af16b8a90>

|  |  |
| --- | --- |
|  | 1600 |
| 1400 |
| 1200 |
| Ill | 1000 |



*2*



**SCATTER PLOT**

l)Customers paying high monthly charges for short tenures are disconnecting

8)() 2)Customers paying high

6 monthly charges for long

!mo tenures continuing with



their subscription plans, as it is reasonable cost

ax>

0



0 10 ro :ll «> !i) 00 10

•

tenure

ACCURACY OF VARIOUS MOD ELS

MODELS

ACCURACY

KNN 62%

SVM 76%

SVM MOD EL

|  |  |  |
| --- | --- | --- |
| H | *#5V1*  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 = | 0.3,random\_state = 0) |
|  | classifier = SVC{kernel = 'rbf') classifier.fit(x\_train, y\_train) y\_pred = classifier.predict(x\_test) |  |

ac.(uracyl = rnetrics.acc1Jracy\_score(y\_testJ yjp red )

##### print("Accuracy:",accuracyl)

Accuracy : 0.7597955706984668

print("Precision :11 ,metrics .precision\_score(y\_testJ y\_pred))

##### print("Recall:",metrics.recall\_score(y\_test, y\_pred) )

Precision: 0.7796610169491526

Recall: 0.10143329658213891

H 1 *#p l ot:t ing t:he roe sea tar*

plt.title(" Receiver Operating Characteristic')

plt.plot(fpr ... i::p r... ·b'... label = • AUC = 0 .2·f ' % roc\_auc) plt.legend(loc = " lower right " )

plt.plot({0, 1], [0... 1] , ·r- - ')

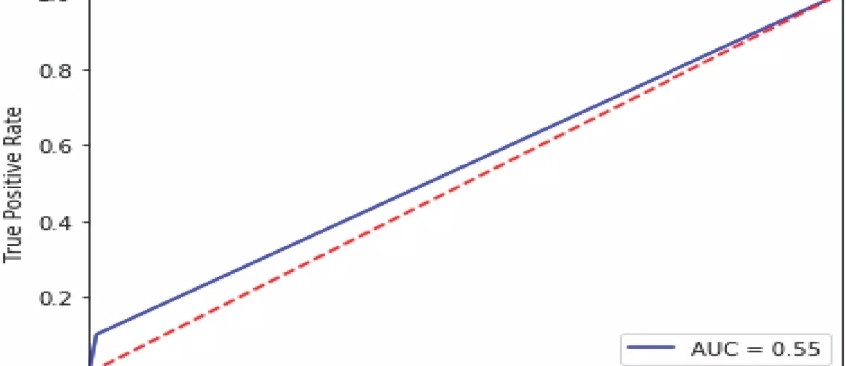
plt.xlim( [e, 1])

plt.ylim([0, 1.] )

p lt.ylabel('True Positive Rate" ) plt.xlabel(" False Positive Rate') p1't .show

c )

o\_ o 0-4



io .---

Receiver Ope rating C haracteri1stic

----,,...,

0..0 --

--.-

----..-

---.

.--

---t

Fals.e Positive Rate

M ETR ICS EVALUATION:

*CONFUSION MATRIX:*



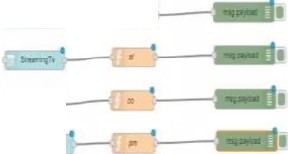
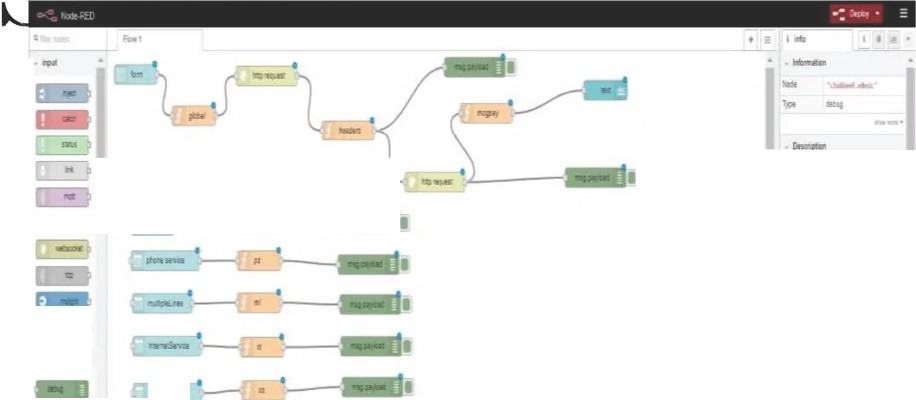
815

92

*ACCURACY: 76.12% PRECISION : 77.96%*

# •

l"J UlJ t\, lJ ltllVl WA:l' U l"J



Q'11T Tn TI\'\

·-

I

..

•

.:.., lilf«lE.""ftSICI ;e:r\_e•n

11

""2'lll-

!"1

'l-=--411 *:........ ..,,*

\_,

:.t:e1rcea

l!J9''-"'! rt\l:J!"f'

"!S.t.1.S,...,oia!l:l!SW

•

·-

·li<IMI

?it-.;1.:tCSJDi •r.on:

ft'tiht'!!f"l!.Cil."'; 1

aw:

•

..

•-

w-tn:SP..

. ...:.,.:r:a

l't'r.r:it:e-....:...

:C:C!:.r2: '! !1!11itl'Sl!

•-

Is=;; *i*:m»!

..,

.....,....

l:'l!llr"llCL!ll "

'"\*""'9•T.Hr: ""'. 1:•'

r.1n-•:r..•:f

7,.mw-.c.t!IMl 'GI

E1J

- ·

--..



telchurn

text 0

tgender Ma1e-

prone service No m.1ttiptelines No

tntemetServ ice No Onlir-eSecurity No Teoh Support No



Contract Two y · ear

5ank tra n :sfer ( il u tomi!t *CJ*

NODE RED FLOW WIRING

PaymentMett-od

tenure ,..

Montt>lyCharges •



Tota Charges •

CANCEL

UI OF MODEL •

FINDINGS AND SUGG ESTIONS

)-Try to offer the better service for the churn customers ,see how much this impact before and later .Some may use your service better move them to your active customers.

)-Take the feedback and suggestions with in period of time and improve it ,strive for better communication.

);> When your are taking the any change in plans of your business just predict the positive and negative share of that plan. If it is negative prepare the solution before so You can handy easily.

**HOW TO REDUCE CUSTOMER CHURN**

Lean into your best customers.

Be proactive with communication.

Define a roadmap for your new customers. Offer incentives.

Ask for feedback often.

Analyze churn when it happens. Stay competitive.