INTELLIGENT CUSTOMER RETENTION: USING MACHINE LEARNING FOR ENHANCED PREDICTION OF TELECOM CUSTOMER CHURN

Table Of Content

1.INTRODUCTION

1.1 Overview

We are introducing our project Intelligent Customer Retention: Using Machine Learning For Enhanced Prediction Of Telecom Customer Churn

- Customer churn is often referred to as customer attrition, or customer defection which is the rate at which the customers are lost. Customer churn is a major problem and one of the most important concerns for large companies. Due to the direct effect on the revenues of the companies, especially in the telecom field, companies are seeking to develop means to predict potential customer to churn. Looking at churn, different reasons trigger customers to terminate their contracts, for example better price offers, more interesting packages, bad service experiences or change of customers' personal situations.

- Customer churn has become highly important for companies because of increasing competition among companies, increased importance of marketing strategies and conscious behaviour of customers in the recent years. Customers can easily trend toward alternative services. Companies must develop various strategies to prevent these possible trends, depending on the services they provide. 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.

- Telecommunication industry always suffers from a very high churn rates when one industry offers a better plan than the previous there is a high possibility of the customer churning from the present due to a better plan in such a scenario it is very difficult to avoid losses but through prediction we can keep it to a minimal level.

- Telecom companies often use customer churn as a key business metrics to predict the number of customers that will leave a telecom service provider. A machine learning model can be used to identity the probable churn customers and then makes the necessary business decisions.

1.2 Purpose

Intelligent customer retention using machine learning and telecom customer churn are important applications of machine learning in the telecommunications industry. Customer churn is a major challenge for telecom companies, as losing a customer not only means losing revenue but also the potential for future revenue. Therefore, it is crucial for telecom companies to identify and retain customers who are at risk of churning.

Machine learning can be used to analyze large volumes of customer data and identify patterns that are indicative of churn. By analyzing customer behavior, usage patterns, and other data points, machine learning algorithms can predict which customers are likely to churn and why. This information can then be used to develop targeted retention strategies, such as offering incentives or personalized promotions to keep customers engaged.

The benefits of using machine learning for customer retention in the telecom industry include:

- **Increased customer satisfaction**: By identifying and addressing issues before they become major problems, machine learning can help improve customer satisfaction and reduce churn rates.

Better decision-making: Machine learning can provide telecom companies with valuable insights into customer behavior and preferences, allowing them to make more informed decisions about product development, pricing, and other business strategies.

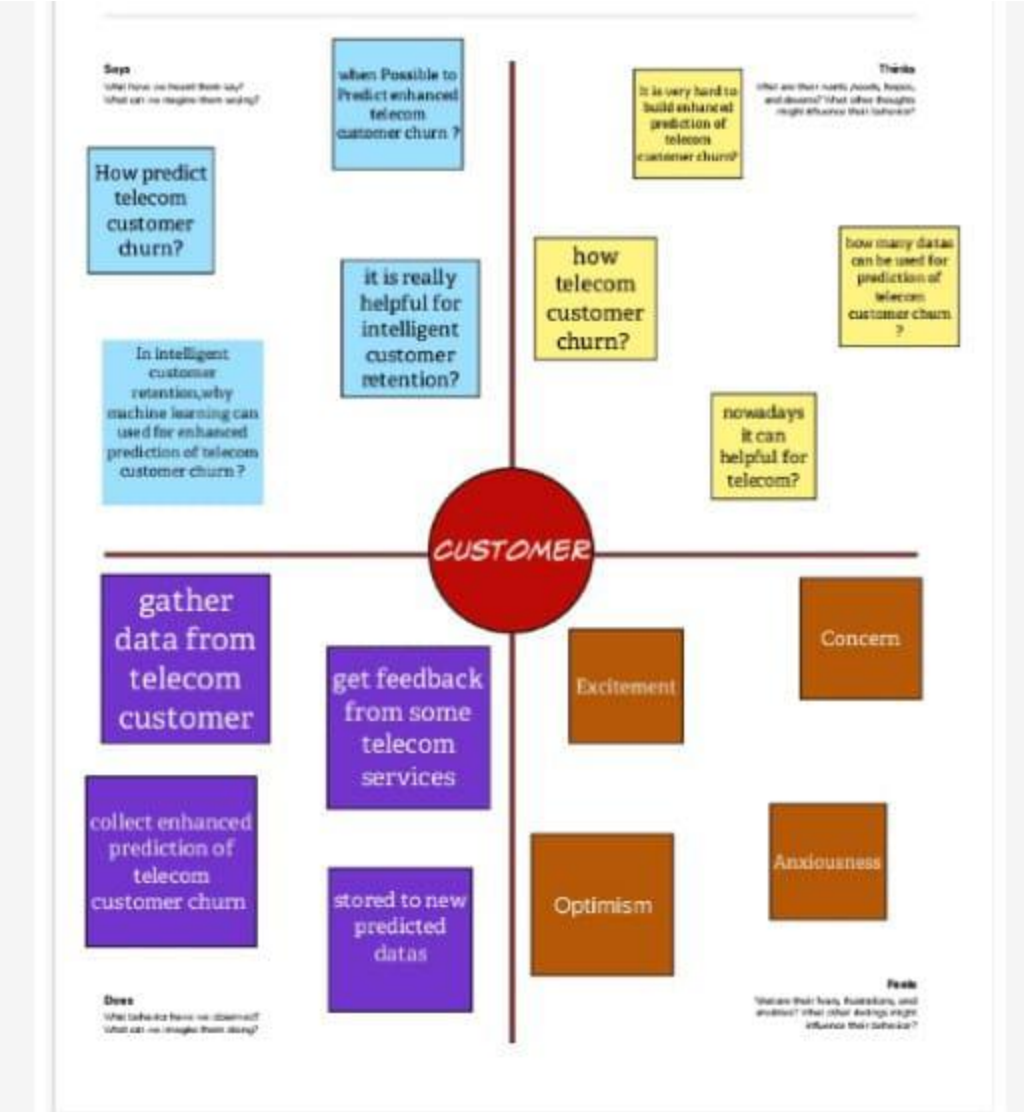
- **Cost savings**: Retaining existing customers is often less expensive than acquiring new ones, so reducing churn rates can lead to significant cost savings for telecom companies.

- **Competitive advantage**: Telecom companies that are able to effectively retain customers through targeted retention strategies are better positioned to compete in a crowded marketplace.

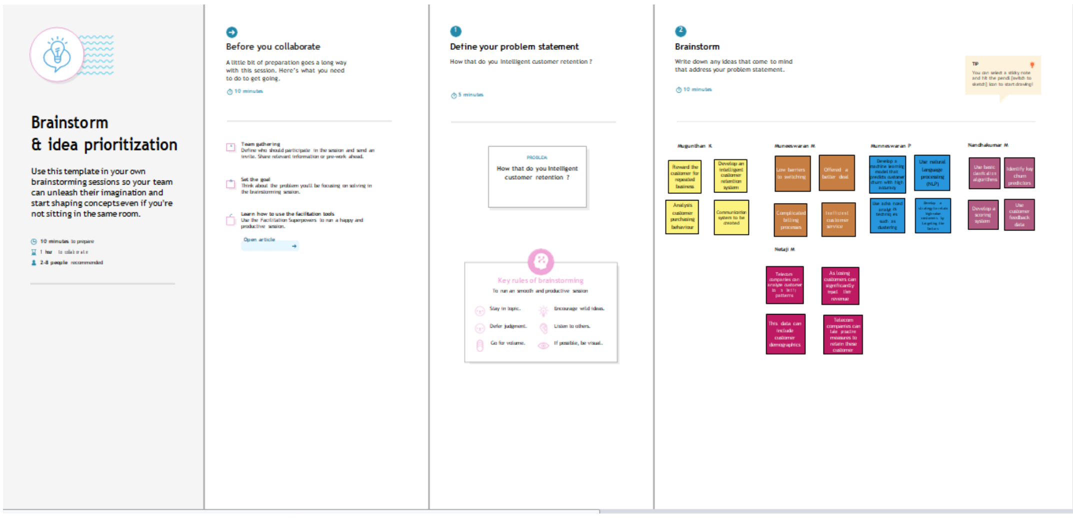
Overall, intelligent customer retention using machine learning and telecom customer churn analysis can help telecom companies increase customer satisfaction, reduce churn rates, and gain a competitive advantage in the market.

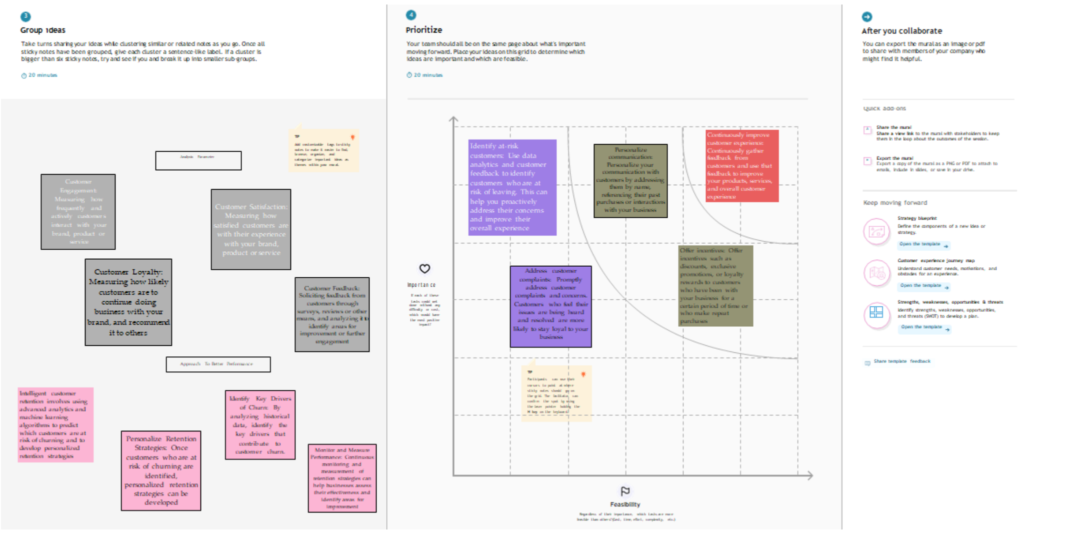
2. PROBLEM DEFINITION AND DESIGN THINKING

2.1.Empathy Map



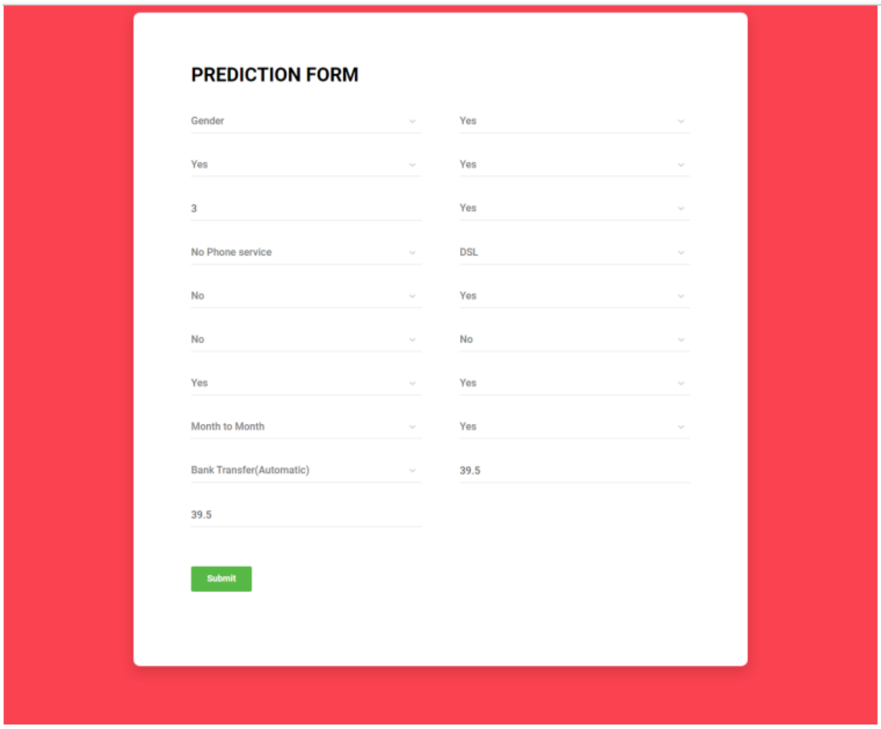
2.2 Ideation & Brainstorming Map



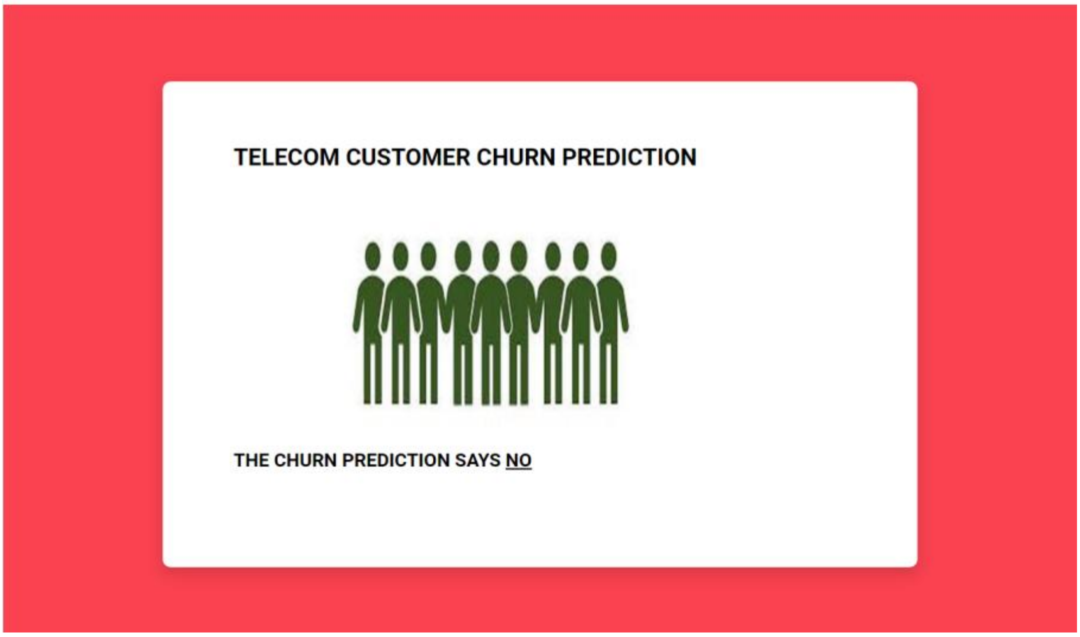


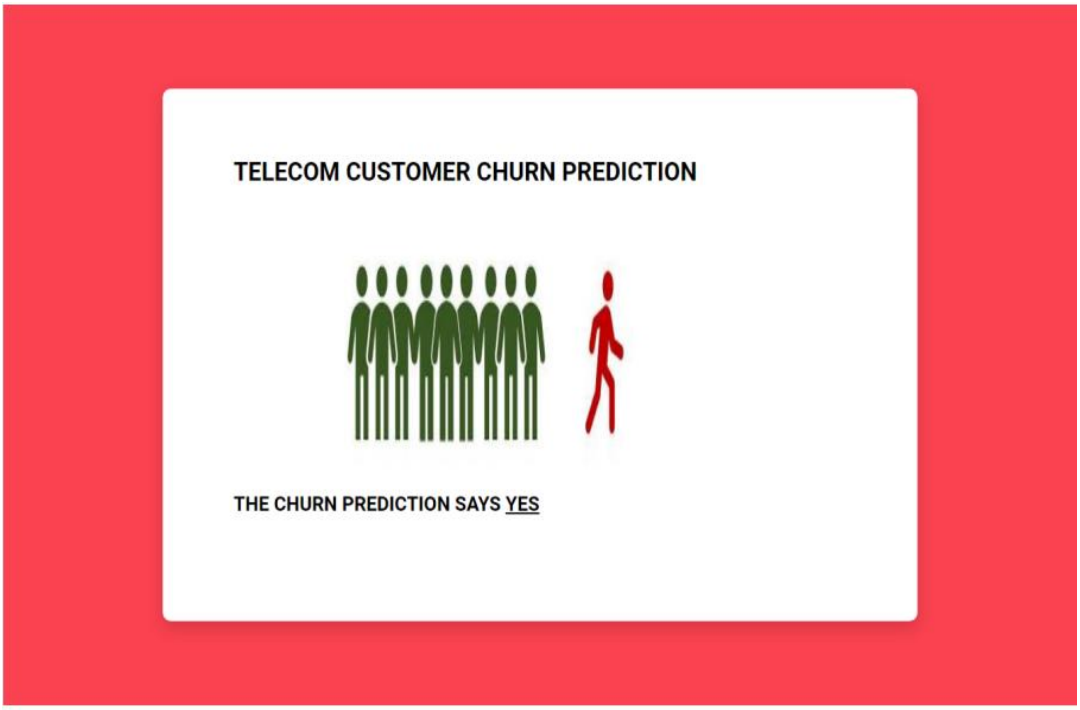
3.RESULT

TELECOM CUSTOMER CHURN PREDICTION
Customer churn has become highly important for companies because of increasing competition among companies, increased importance of marketing strategies and conscious behaviour of customers in the recent years. Customers can easily trend toward alternative services. Companies must develop various strategies to prevent these possible trends, depending on the services they provide. 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.
EXIT
Click me to continue with prediction



PREDICTION FORM
Gender
Yes
Yes
Yes
3
Yes
No Phone service
DSL
No
Yes
No
No
Yes
Yes
Month to Month
Yes
Bank Transfer(Automatic)
39.5
39.5
Submit


TELECOM CUSTOMER CHURN PREDICTION
THE CHURN PREDICTION SAYS NO


TELECOM CUSTOMER CHURN PREDICTION
THE CHURN PREDICTION SAYS YES

4.ADVANTAGES AND DISADVANTAGES

Advantages:

1. ***Personalization:*** Machine learning algorithms can analyze vast amounts of data on customer behavior and preferences to create personalized retention strategies. This can increase the likelihood of retaining customers by providing tailored offers and recommendations.

2. ***Improved accuracy:*** Machine learning algorithms can identify patterns in customer behavior that might go unnoticed by humans. This can help companies to better understand the factors that contribute to customer churn and take proactive steps to address them.

3. ***Cost-effective:*** Implementing machine learning algorithms can reduce the cost of customer retention strategies by automating many of the tasks that would otherwise be performed manually.

4. ***Real-time decision making:*** Machine learning algorithms can make quick decisions based on real-time data, enabling companies to respond to customer churn in a timely manner.

Disadvantages:

1. ***Complexity:*** Implementing machine learning algorithms can be complex and require specialized expertise. This can be a significant barrier to entry for some companies.

2. ***Data quality:*** Machine learning algorithms are only as good as the data they are trained on. If the data is incomplete, inaccurate, or biased, the algorithms may produce inaccurate results.

3. ***Privacy concerns:*** Machine learning algorithms require access to large amounts of customer data, which can raise privacy concerns among customers.

4. ***Unforeseen outcomes:*** Machine learning algorithms can produce unexpected outcomes that may not align with a company's goals or values. It is important to monitor these outcomes and adjust the algorithms accordingly.

5.APPLICATIONS

Intelligent customer retention using machine learning can be applied in various industries and businesses. Here are some of the application areas:

1. ***E-commerce:*** E-commerce companies can use machine learning algorithms to analyze customer behavior, purchase history, and preferences to create personalized retention strategies. This can include recommendations for products or services, special offers, and targeted marketing campaigns.

2. ***Telecommunications:*** Telecom companies can use machine learning algorithms to identify the factors that contribute to customer churn and create proactive retention strategies. This can include personalized offers, targeted marketing campaigns, and improved customer service.

3. ***Banking and finance:*** Banks and financial institutions can use machine learning algorithms to analyze customer data and create personalized retention strategies. This can include personalized financial advice, targeted marketing campaigns, and improved customer service.

4. ***Healthcare:*** Healthcare providers can use machine learning algorithms to analyze patient data and create personalized retention strategies. This can include personalized health recommendations, targeted marketing campaigns, and improved patient communication.

5. ***Retail:*** Retail companies can use machine learning algorithms to analyze customer data and create personalized retention strategies. This can include personalized recommendations, targeted marketing campaigns, and improved customer service.

6. ***Hospitality:*** Hospitality companies can use machine learning algorithms to analyze customer data and create personalized retention strategies. This can include personalized recommendations for travel and

accommodation, targeted marketing campaigns, and improved customer service.

6.CONCLUSION

Intelligent customer retention using machine learning can be a powerful tool for telecom companies to reduce customer churn. By analyzing customer behavior and predicting their likelihood to churn, telecom companies can proactively engage with customers and offer them personalized incentives to stay loyal.

- Machine learning algorithms can analyze vast amounts of customer data, including call logs, text messages, and social media interactions, to identify patterns and make accurate predictions about customer behavior. By using this information, telecom companies can create targeted retention campaigns that are tailored to each individual customer's needs and preferences.

- By using machine learning to optimize their retention strategies, telecom companies can reduce customer churn and increase customer satisfaction. Additionally, they can gain insights into customer behavior and preferences, which can be used to improve products and services and better serve their customers.

- Overall, intelligent customer retention using machine learning is an effective and efficient way for telecom companies to reduce churn and increase customer loyalty. It is a rapidly evolving field, and with continued investment in data analytics and machine learning technology, it is likely to become even more powerful in the future.

7.FUTURE SCOPE

The future scope of intelligent customer retention and telecom customer churn using machine learning is vast and promising. As machine learning technology advances, it will become even more effective in predicting and preventing customer churn. Some potential future developments include:

1. ***Real-time intervention:*** As machine learning algorithms become faster and more sophisticated, telecom companies will be able to intervene in real-time to prevent customers from leaving. This will enable companies to offer targeted incentives and personalized solutions to keep customers happy and loyal.

2. ***Enhanced customer experience:*** Machine learning algorithms will enable telecom companies to better understand their customers' preferences, needs, and behaviors. This will allow companies to tailor their products and services to meet these needs, providing a better customer experience and reducing churn.

3. ***Increased automation:*** As machine learning technology continues to improve, telecom companies will be able to automate many of their retention strategies. This will reduce the need for human intervention, making retention campaigns more efficient and cost-effective.

4. ***Integration with other technologies:*** Machine learning will be integrated with other technologies such as natural language processing and voice recognition, enabling companies to interactwith customers more effectively and provide better customer service.

5.***Advanced predictive models:*** Machine learning algorithms will continue to improve, allowing companies to predict customer behavior more accurately. This will enable telecom companies to identify and address potential issues before they become serious enough to cause customers to churn.

8.APPENDIX

A.Source Code

```
[ ] import numpy as np
    import pandas as pd
    import matplotlib.pyplot as plt
    import seaborn as sns
```

```
data = pd.read_csv("WA_Fn-UseC_-Telco-Customer-Churn.csv")
```

```
[ ] data.head()
```

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	No	No	No	Month-to-month	Yes	Electronic check	29.85	29.85	No
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	No	No	No	One year	No	Mailed check	56.95	1889.5	No
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	No	No	No	Month-to-month	Yes	Mailed check	53.85	108.15	Yes
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	Yes	No	No	One year	No	Bank transfer (automatic)	42.30	1840.75	No
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	No	No	No	Month-to-month	Yes	Electronic check	70.70	151.65	Yes

5 rows × 21 columns

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7043 entries, 0 to 7042
Data columns (total 21 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   customerID        7043 non-null   object
 1   gender            7043 non-null   object
 2   SeniorCitizen     7043 non-null   int64
 3   Partner           7043 non-null   object
 4   Dependents        7043 non-null   object
 5   tenure            7043 non-null   int64
 6   PhoneService      7043 non-null   object
 7   MultipleLines     7043 non-null   object
 8   InternetService   7043 non-null   object
 9   OnlineSecurity    7043 non-null   object
 10  OnlineBackup      7043 non-null   object
 11  DeviceProtection  7043 non-null   object
 12  TechSupport       7043 non-null   object
 13  StreamingTV       7043 non-null   object
 14  StreamingMovies   7043 non-null   object
 15  Contract          7043 non-null   object
 16  PaperlessBilling  7043 non-null   object
 17  PaymentMethod     7043 non-null   object
 18  MonthlyCharges    7043 non-null   float64
 19  TotalCharges      7043 non-null   object
 20  Churn             7043 non-null   object
dtypes: float64(1), int64(2), object(18)
memory usage: 1.1+ MB
```

```
data['TotalCharges']=pd.to_numeric(data['TotalCharges'],errors='coerce')
```

```
[ ] data.head()
```

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	No	No	No	Month-to-month	Yes	Electronic check	29.85	29.85	No
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	No	No	No	One year	No	Mailed check	56.95	1889.50	No
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	No	No	No	Month-to-month	Yes	Mailed check	53.85	108.15	Yes
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	Yes	No	No	One year	No	Bank transfer (automatic)	42.30	1840.75	No
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	No	No	No	Month-to-month	Yes	Electronic check	70.70	151.65	Yes

5 rows × 21 columns

```
[ ] data.describe()
```

	SeniorCitizen	tenure	MonthlyCharges	TotalCharges
count	7043.000000	7043.000000	7043.000000	7032.000000
mean	0.162147	32.371149	64.761692	2283.300441
std	0.368612	24.559481	30.090047	2266.771362
min	0.000000	0.000000	18.250000	18.800000
25%	0.000000	9.000000	35.500000	401.450000
50%	0.000000	29.000000	70.350000	1397.475000
75%	0.000000	55.000000	89.850000	3794.737500
max	1.000000	72.000000	118.750000	8684.800000

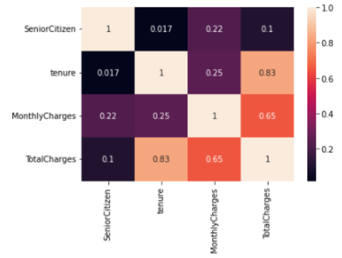
```
[ ] data.isnull().any()

    customerID          False
    gender              False
    SeniorCitizen       False
    Partner             False
    Dependents          False
    tenure              False
    PhoneService        False
    MultipleLines       False
    InternetService     False
    OnlineSecurity      False
    OnlineBackup        False
    DeviceProtection    False
    TechSupport         False
    StreamingTV         False
    StreamingMovies     False
    Contract            False
    PaperlessBilling    False
    PaymentMethod       False
    MonthlyCharges      False
    TotalCharges         True
    Churn               False
    dtype: bool
```

```
[ ] data['TotalCharges'].fillna(data['TotalCharges'].median(),inplace=True)
```

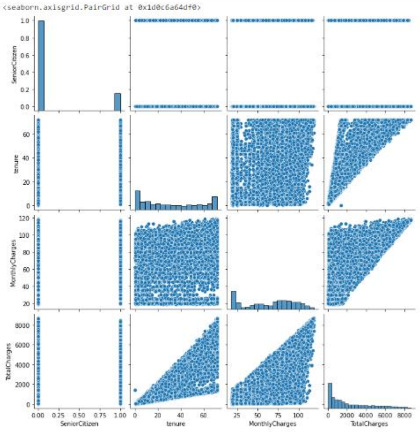
```
sns.heatmap(data.corr(),annot=True)

<AxesSubplot:>
```




```
[ ] sns.pairplot(data=data,markers=["^","v"],palette="inferno")

<seaborn.axisgrid.PairGrid at 0x1d0c6a64df0>
```




```
[ ] from sklearn.preprocessing import OneHotEncoder
    one=OneHotEncoder()
    a= one.fit_transform(x[:,6:7]).toarray()
    b= one.fit_transform(x[:,7:8]).toarray()
    c= one.fit_transform(x[:,8:9]).toarray()
    d= one.fit_transform(x[:,9:10]).toarray()
    e= one.fit_transform(x[:,10:11]).toarray()
    f= one.fit_transform(x[:,11:12]).toarray()
    g= one.fit_transform(x[:,12:13]).toarray()
    h= one.fit_transform(x[:,13:14]).toarray()
    i= one.fit_transform(x[:,14:15]).toarray()
    j= one.fit_transform(x[:,16:17]).toarray()
    x=np.delete(x,[6,7,8,9,10,11,12,13,14,16], axis=1)
    x=np.concatenate((a,b,c,d,e,f,g,h,i,j,x), axis=1)
```

```
x_train
```

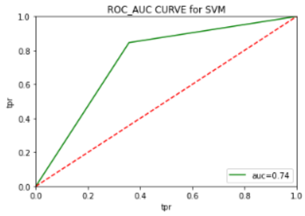
```
array([[ 0.32472445, -0.32287108, -0.13749918, ...,  0.77678217,
         0.99770032,  0.39690084],
       [ 1.09105968, -0.32287108, -0.90679393, ..., -1.3967162 ,
         0.4217374 ,  0.46094636],
       [-0.98894352, -0.32287108,  1.18124205, ..., -1.3967162 ,
        -1.54177849, -0.13058306],
       ...,
       [ 1.09105968, -0.32287108, -0.90679393, ...,  0.77678217,
         0.79873246, -0.49650017],
       [ 1.09105968, -0.32287108, -0.90679393, ...,  0.77678217,
        -0.7847252 , -0.91646012],
       [ 1.09105968, -0.32287108, -0.90679393, ..., -1.3967162 ,
        -1.66918886, -0.89416759]])
```

```
x_test
```

```
array([[-0.98063706, -0.32775742,  1.17624977, ...,  0.76490571,
        -0.28119439, -0.9132475 ],
       [ 1.09792446, -0.32775742, -0.90847981, ...,  0.76490571,
        -1.71347878, -0.77761264],
       [-0.98063706, -0.32775742,  1.17624977, ...,  0.76490571,
         0.47887458,  0.30898783],
       ...,
       [-0.98063706, -0.32775742,  1.17624977, ..., -1.40569096,
         0.43582222,  0.86159919],
       [ 1.09792446, -0.32775742, -0.90847981, ...,  0.76490571,
        -1.68025118, -0.85047278],
       [ 1.09792446, -0.32775742, -0.90847981, ...,  0.76490571,
         0.10825841, -0.87643352]])
```

```python
import matplotlib.pyplot as plt
plt.title("ROC_AUC CURVE for SVM")
plt.plot(fpr,tpr,'g',label='auc=%0.2f'%roc_auc)
plt.plot([0,1],[0,1],'r--')
plt.xlim([0,1])
plt.ylim([0,1])
plt.xlabel('tpr')
plt.ylabel('tpr')
plt.legend(loc='lower right')
```

```
<matplotlib.legend.Legend at 0x1d0c9e71cd0>
```




```python
import pickle
pickle.dump(svm,open('churnnew.pkl','wb'))
```