Intelligent Customer Retention: Using Machine Learning for Enhanced Prediction of Telecom Customer Churn

Project Record Template

CHAPTER-1

1.INTRODUCTION

1.1 OVERVIEW

Intelligent Customer Retention is a process that involves using machine learning techniques to predict which customers are likely to churn, and then taking targeted actions to retain them. In the context of telecom, customer churn refers to the loss of customers who switch to another service provider.

The use of machine learning algorithms for customer churn prediction has become increasingly popular in recent years, due to the vast amounts of data generated by telecom companies, as well as the complexity of the factors that influence customer behavior. Machine learning models can analyze this data and identify patterns that are not immediately apparent to humans, allowing companies to make more accurate predictions about which customers are likely to churn.

Once these customers are identified, companies can take targeted actions to retain them, such as offering personalized discounts or improving the quality of customer service. By using machine learning to optimize these retention strategies, companies can reduce their overall churn rate and improve customer satisfaction.

Overall, the use of machine learning for intelligent customer retention represents a significant opportunity for telecom companies to improve their bottom line and better serve their customers.

Intelligent customer retention refers to the use of machine learning techniques to predict and prevent customer churn in the telecom industry. Customer churn is a critical issue for telecom companies as it results in significant revenue loss and reduces customer loyalty.

By using machine learning algorithms, telecom companies can analyze customer data and behavior patterns to identify customers who are at risk of churning. These algorithms can identify patterns in customer behavior, such as changes in usage patterns, complaints, and inquiries. By identifying these patterns, telecom companies can take proactive measures to retain their customers, such as targeted marketing campaigns or offering incentives.

Intelligent customer retention also involves the use of predictive analytics to forecast the likelihood of a customer churning. Predictive analytics algorithms use historical customer data, such as purchase history, call history, and demographics, to build models that can predict the likelihood of a customer leaving.

The ultimate goal of intelligent customer retention is to prevent customer churn and improve customer satisfaction. By using machine learning and predictive analytics, telecom companies can proactively address customer issues and provide personalized solutions, thereby improving customer loyalty and reducing revenue loss.

1.1 PURPOSE

The purpose of intelligent customer retention using machine learning for enhanced prediction of telecom customer churn is to reduce the rate of customer churn, improve customer satisfaction, and increase revenue for telecom companies. The telecom industry is highly competitive, and retaining customers is crucial for businesses to remain profitable.

By using machine learning algorithms to analyze customer behavior and predict the likelihood of churn, telecom companies can take proactive measures to prevent customers from leaving. This could involve providing targeted marketing campaigns or personalized incentives to keep customers engaged and loyal.

The use of machine learning and predictive analytics can also help telecom companies to identify and address issues that may lead to customer churn. By analyzing customer data, companies can identify patterns and trends in customer behavior and take proactive measures to address any issues before they become major problems.

Overall, the purpose of intelligent customer retention is to help telecom companies to retain their customers and improve customer satisfaction, ultimately leading to increased revenue and long-term business success.

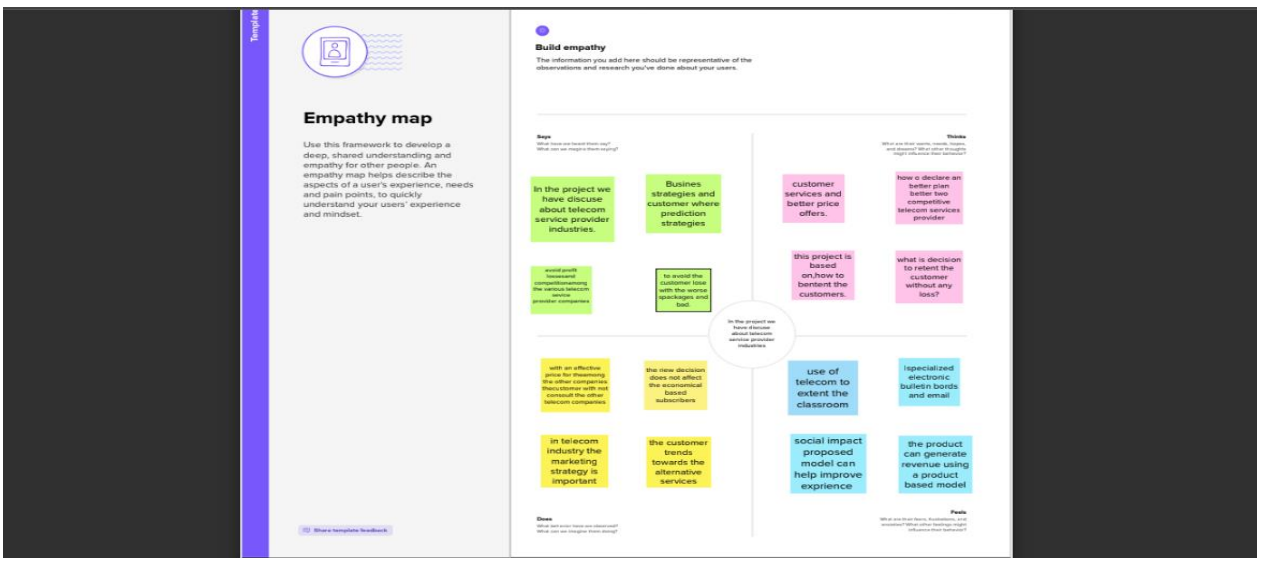
The purpose of Intelligent Customer Retention is to reduce customer churn in the telecom industry through the use of machine learning techniques. Customer churn, which refers to the rate at which customers cancel their subscriptions or switch to other providers, is a major issue in the telecom industry. By using machine learning algorithms and predictive analytics, telecom companies can identify customers who are at risk of churning and take proactive measures to retain them.

The use of machine learning in customer retention allows telecom companies to analyze large amounts of customer data and identify patterns and trends that may indicate a customer is likely to churn. These patterns may include changes in usage patterns, complaints, or inquiries. By identifying these patterns, companies can take proactive measures, such as targeted marketing campaigns or personalized incentives, to retain customers and improve their satisfaction.

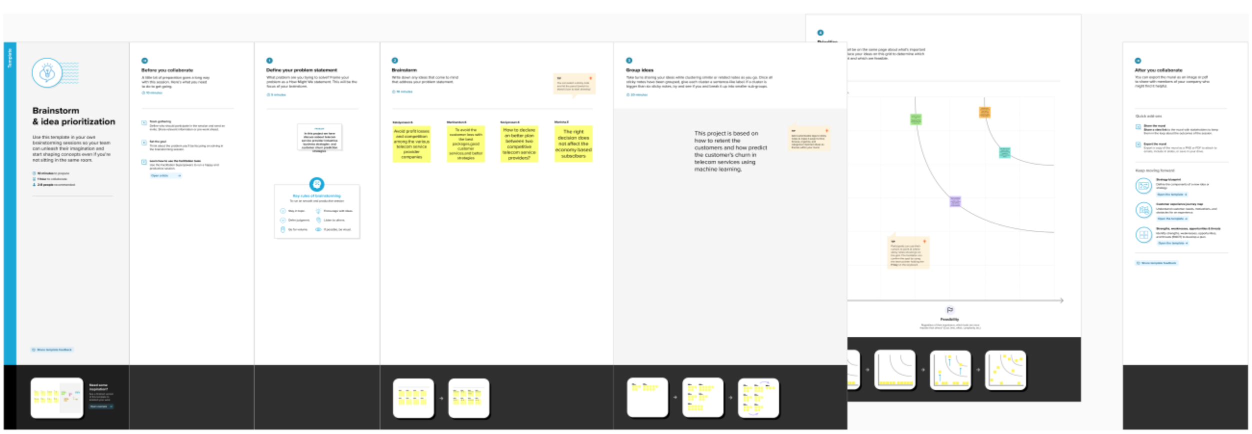
CHAPTER-2

2. PROBLEM DEFINITION & DESIGN THINKING

2.1 PROBLEM DEFINITION :



2.2 IDEATION & BRAINSTORMING MAP



CHAPTER-3

3. RESULT

Result 1:

- Import all the tools we need.
- **A**ll needed tools import sucessful.

Result 2:

	Tenure	Monthly_Charges	Yearly_Charge	Admin_Tickets	Tech_Tickets
count	1142.000000	1142.000000	1142.000000	1142.000000	1142.000000
mean	33.295972	798.203590	9578.443082	0.513135	0.684764
std	24.188530	237.640267	2851.683204	1.296967	1.550357
min	1.000000	189.500000	2274.000000	0.000000	0.000000
25%	10.000000	701.500000	8418.000000	0.000000	0.000000
50%	31.000000	848.500000	10182.000000	0.000000	0.000000
75%	56.000000	980.750000	11769.000000	0.000000	0.000000
max	72.000000	1174.500000	14094.000000	5.000000	9.000000

Result 3:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1142 entries, 0 to 1141
Data columns (total 22 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   Gender             1142 non-null   object
 1   Senior_Citizen     1142 non-null   object
 2   Partner            1142 non-null   object
 3   Dependents         1142 non-null   object
 4   Tenure             1142 non-null   int64
 5   Phone_Service      1142 non-null   object
 6   Multiple_Lines     1142 non-null   object
 7   Internet_Service   1142 non-null   object
 8   Online_Security    1142 non-null   object
 9   Online_Backup      1142 non-null   object
 10  Device_Protection  1142 non-null   object
 11  Tech_Support       1142 non-null   object
 12  Streaming_TV       1142 non-null   object
 13  Streaming_Movies   1142 non-null   object
 14  Contract           1142 non-null   object
 15  Paper_less_Billing 1142 non-null   object
 16  Payment_Method     1142 non-null   object
 17  Monthly_Charges    1142 non-null   float64
 18  Yearly_Charge      1142 non-null   int64
 19  Admin_Tickets      1142 non-null   int64
 20  Tech_Tickets       1142 non-null   int64
 21  Churn              1142 non-null   object
dtypes: float64(1), int64(4), object(17)
memory usage: 196.4+ KB
```

Result 4:

```
Gender               False
Senior_Citizen       False
Partner              False
Dependents           False
Tenure               False
Phone_Service        False
Multiple_Lines       False
Internet_Service     False
Online_Security      False
Online_Backup        False
Device_Protection    False
Tech_Support         False
Streaming_TV         False
Streaming_Movies     False
Contract             False
Paper_less_Billing   False
Payment_Method       False
Monthly_Charges      False
Yearly_Charge        False
Admin_Tickets        False
Tech_Tickets         False
Churn                False
dtype: bool
```

Result 5:

```
Gender               0
Senior_Citizen       0
Partner              0
Dependents           0
Tenure               0
Phone_Service        0
Multiple_Lines       0
Internet_Service     0
Online_Security      0
Online_Backup        0
Device_Protection    0
Tech_Support         0
Streaming_TV         0
Streaming_Movies     0
Contract             0
Paper_less_Billing   0
Payment_Method       0
Monthly_Charges      0
Yearly_Charge        0
Admin_Tickets        0
Tech_Tickets         0
Churn                0
dtype: int64
```

Result 6:

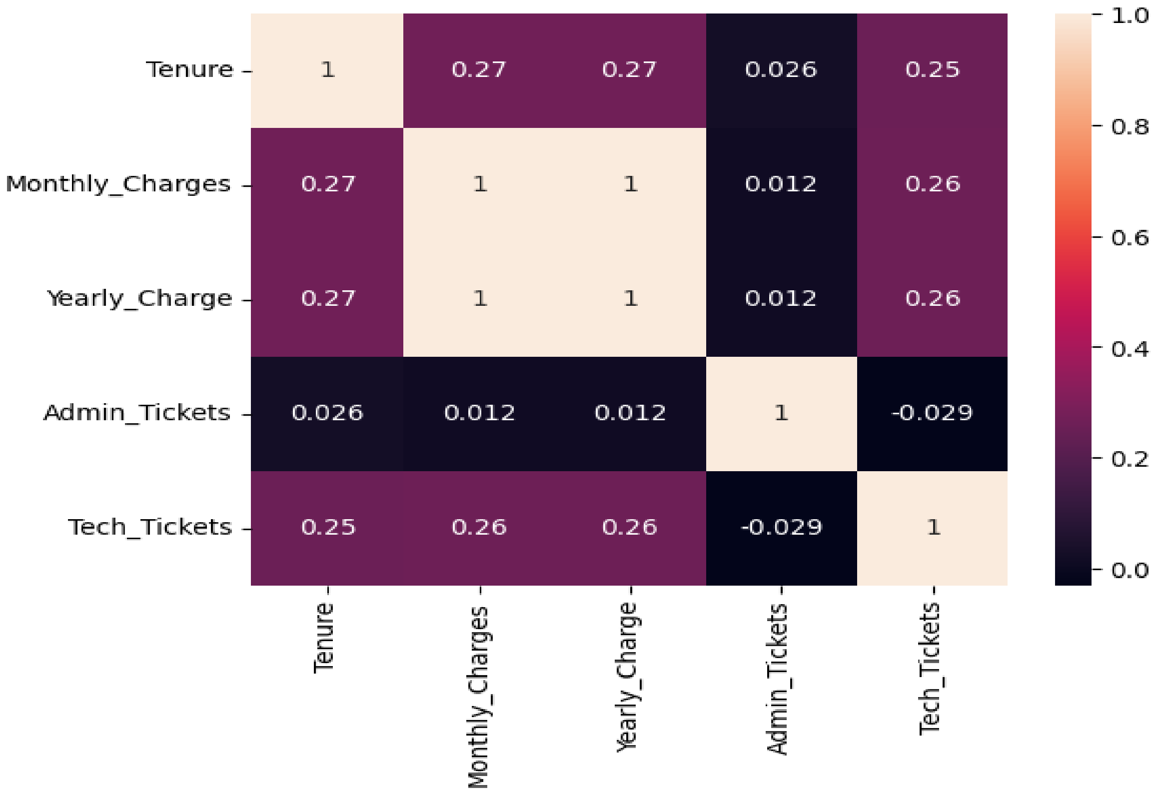
```
type: int64
```

```
df.corr()
```

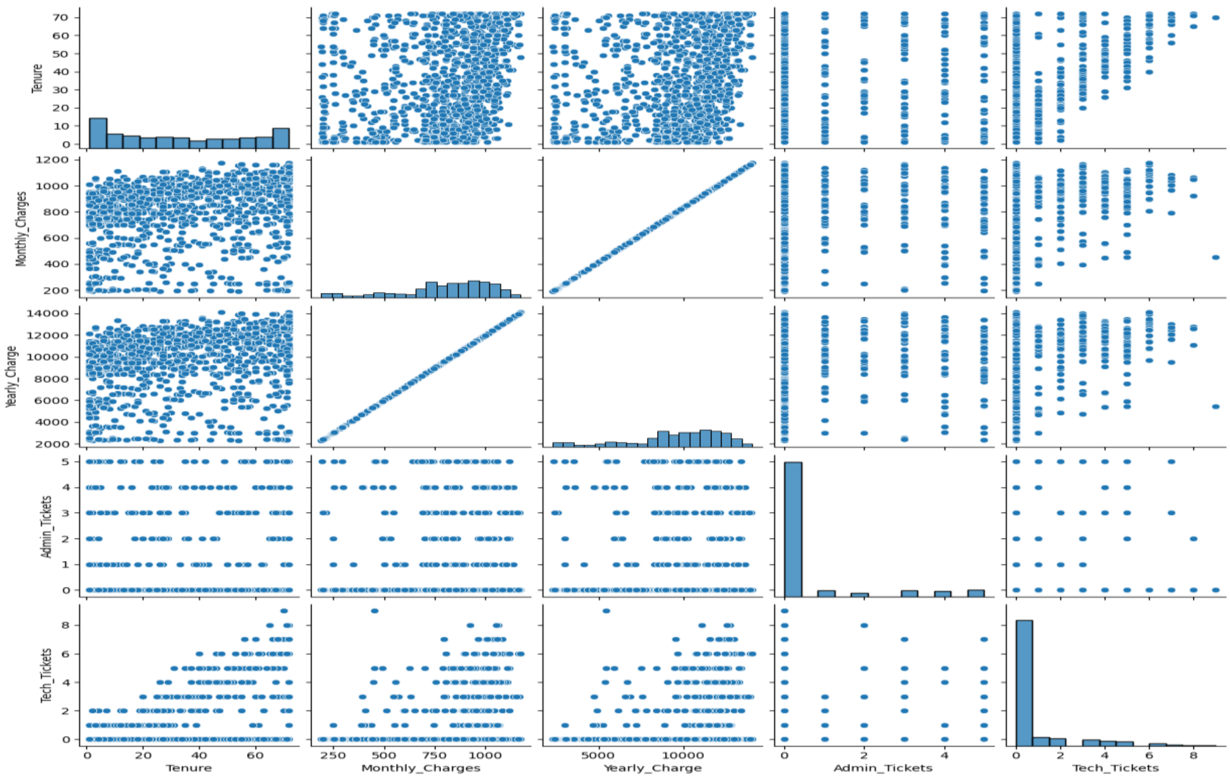
```
<ipython-input-12-2f6f6606aa2c>:1: FutureWarning: The default value of numeric_only in
  df.corr()
```

	Tenure	Monthly_Charges	Yearly_Charge	Admin_Tickets	Tech_Tickets
Tenure	1.000000	0.267175	0.267175	0.026220	0.254684
Monthly_Charges	0.267175	1.000000	1.000000	0.012450	0.257708
Yearly_Charge	0.267175	1.000000	1.000000	0.012450	0.257708
Admin_Tickets	0.026220	0.012450	0.012450	1.000000	-0.028886
Tech_Tickets	0.254684	0.257708	0.257708	-0.028886	1.000000

Result 7:



Result 8:



Result 9:

```
['Male' 'Female']
['Yes']
['No' 'Yes']
['No' 'Yes']
[ 1 71  2 43 25  8 60 18 66 64 56 30 37 27 72 58 15  7 23 11 65 13  4 29
  3 57 19 16 20 21 24  6 35 32 38  5 54 28 17 55 42 33 48 52 49 22 61 68
 50 47 44 10 40  9 46 70 26 62 53 45 14 12 59 39 34 69 67 41 31 36 51 63]
['No' 'Yes']
['No phone service' 'Yes' 'No']
['DSL' 'Fiber optic' 'No']
['No' 'Yes' 'No internet service']
['No' 'Yes' 'No internet service']
['Yes' 'No' 'No internet service']
['No' 'Yes' 'No internet service']
['No' 'Yes' 'No internet service']
['Yes' 'No' 'No internet service']
['Month-to-month' 'Two year' 'One year']
['Yes' 'No']
['Electronic check' 'Credit card (automatic)' 'Bank transfer (automatic)'
 'Mailed check']
[ 396.5  963.5  955.   452.5  902.5  695.   806.5  748.5  954.5 1084.5
 1116.  1105.   747.5  747.   985.   455.5  765.   780.5  564.5  704.5
  880.5  239.5  453.   744.5  764.5  868.   950.   944.   548.   753.5
  744.   402.   450.   411.5 1069.   898.5 1050.   944.5 1070.5  649.5
  745.   761.   893.5 1008.   749.  1011.5  693.5  931.5  824.5  829. ]
```

Result 10:

```
[1 0]
[0]
[0 1]
[0 1]
[ 1 71  2 43 25  8 60 18 66 64 56 30 37 27 72 58 15  7 23 11 65 13  4 29
  3 57 19 16 20 21 24  6 35 32 38  5 54 28 17 55 42 33 48 52 49 22 61 68
 50 47 44 10 40  9 46 70 26 62 53 45 14 12 59 39 34 69 67 41 31 36 51 63]
[0 1]
[1 2 0]
[0 1 2]
[0 2 1]
[0 2 1]
[2 0 1]
[0 2 1]
[0 2 1]
[2 0 1]
[0 2 1]
[1 0]
[2 1 0 3]
[ 396.5  963.5  955.   452.5  902.5  695.   806.5  748.5  954.5 1084.5
 1116.  1105.   747.5  747.   985.   455.5  765.   780.5  564.5  704.5
  880.5  239.5  453.   744.5  764.5  868.   950.   944.   548.   753.5
  744.   402.   450.   411.5 1069.   898.5 1050.   944.5 1070.5  649.5
  745.   761.   893.5 1008.   749.  1011.5  693.5  931.5  824.5  829.
  703.5  359.   826.5  546.5  721.   970.   891.5  419.   541.   307.5
  648.   199.5  854.5  848.   443.   950.5 1099.   546.   939.5  813.5
  706.5  753.   516.5 1051.  1015.5  697.5 1160.5 1021.   206.5  760.5
  814.5  640.5  284.5  807.   941.   943.   743.5  700.5  752.   516.
 1157.5  999.  1071.5 1019.   596.  1108.5  945.5  892.   857.  1052.5
```

Result 11:

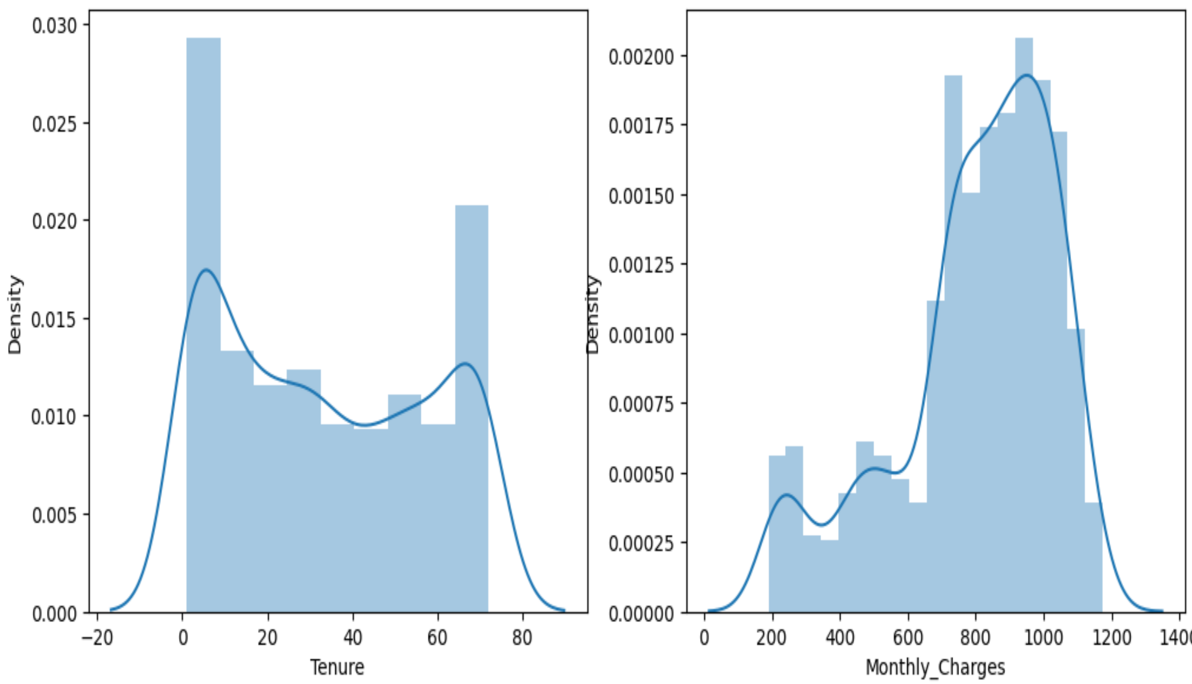
```
RangeIndex: 1142 entries, 0 to 1141
Data columns (total 22 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Gender              1142 non-null   int64
 1   Senior_Citizen      1142 non-null   int64
 2   Partner             1142 non-null   int64
 3   Dependents          1142 non-null   int64
 4   Tenure              1142 non-null   int64
 5   Phone_Service       1142 non-null   int64
 6   Multiple_Lines      1142 non-null   int64
 7   Internet_Service    1142 non-null   int64
 8   Online_Security     1142 non-null   int64
 9   Online_Backup       1142 non-null   int64
 10  Device_Protection   1142 non-null   int64
 11  Tech_Support        1142 non-null   int64
 12  Streaming_TV        1142 non-null   int64
 13  Streaming_Movies    1142 non-null   int64
 14  Contract            1142 non-null   int64
 15  Paper_less_Billing  1142 non-null   int64
 16  Payment_Method      1142 non-null   int64
 17  Monthly_Charges     1142 non-null   float64
 18  Yearly_Charge       1142 non-null   int64
 19  Admin_Tickets       1142 non-null   int64
 20  Tech_Tickets        1142 non-null   int64
 21  Churn               1142 non-null   int64
dtypes: float64(1), int64(21)
memory usage: 196.4 KB
```

Result 12:

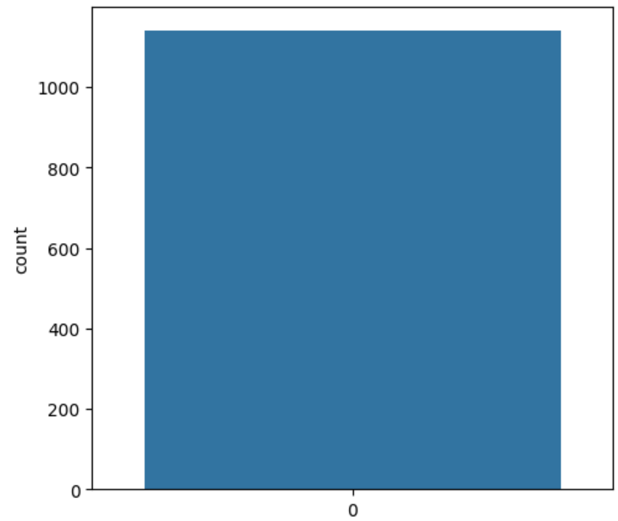
	Gender	Senior_Citizen	Partner	Dependents	Tenure	Phone_Service	Multiple_Lines	Internet_Service	Online_Security	Onlin
Gender	1.000000	NaN	0.129581	0.021089	0.027751	-0.040948	0.023326	-0.003422	-0.025649	
Senior_Citizen	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
Partner	0.129581	NaN	1.000000	0.241486	0.338373	0.007196	0.132502	0.028138	0.138063	
Dependents	0.021089	NaN	0.241486	1.000000	0.140301	-0.008011	0.023701	-0.062808	0.074739	
Tenure	0.027751	NaN	0.338373	0.140301	1.000000	0.026282	0.327828	0.018744	0.355904	
Phone_Service	-0.040948	NaN	0.007196	-0.008011	0.026282	1.000000	0.088121	0.529584	0.035672	
Multiple_Lines	0.023326	NaN	0.132502	0.023701	0.327828	0.088121	1.000000	0.081687	0.027104	
Internet_Service	-0.003422	NaN	0.028138	-0.062808	0.018744	0.529584	0.081687	1.000000	-0.054712	
Online_Security	-0.025649	NaN	0.138063	0.074739	0.355904	0.035672	0.027104	-0.054712	1.000000	
Online_Backup	-0.024636	NaN	0.153988	0.063361	0.442536	-0.020618	0.135011	-0.044324	0.187065	
Device_Protection	-0.006492	NaN	0.176155	0.068843	0.379044	-0.007518	0.142351	-0.014134	0.102174	
Tech_Support	-0.038584	NaN	0.125114	0.105573	0.371972	0.051066	0.049659	-0.047553	0.284943	
Streaming_TV	-0.018202	NaN	0.091361	0.052014	0.307940	0.034052	0.211226	0.107801	0.015490	
Streaming_Movies	0.022925	NaN	0.129274	0.033336	0.315313	0.034741	0.214024	0.112409	0.052287	
Contract	0.004318	NaN	0.221359	0.090539	0.658723	0.028964	0.103082	0.002550	0.384816	
Paper_less_Billing	-0.001248	NaN	0.014360	0.039750	-0.014418	0.005586	0.119575	0.058354	-0.111850	
Payment_Method	-0.021133	NaN	-0.132219	-0.020493	-0.379280	-0.043644	-0.084505	0.020744	-0.199649	
Monthly_Charges	-0.053998	NaN	0.125139	0.030615	0.267175	0.514955	0.382546	0.257728	0.046362	

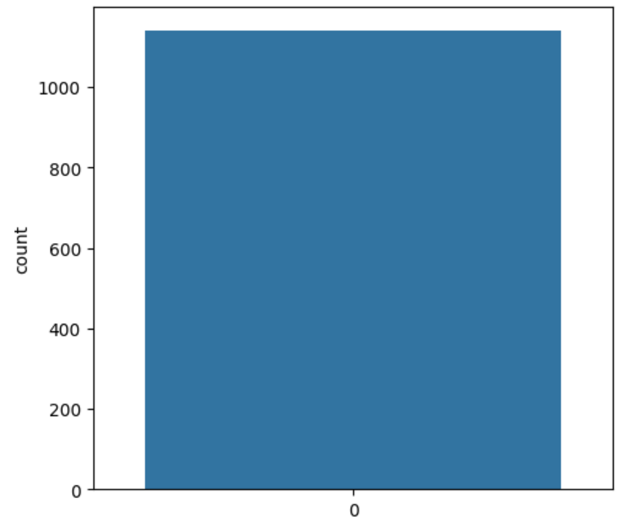
0s completed at 5:04 PM

Result 13:

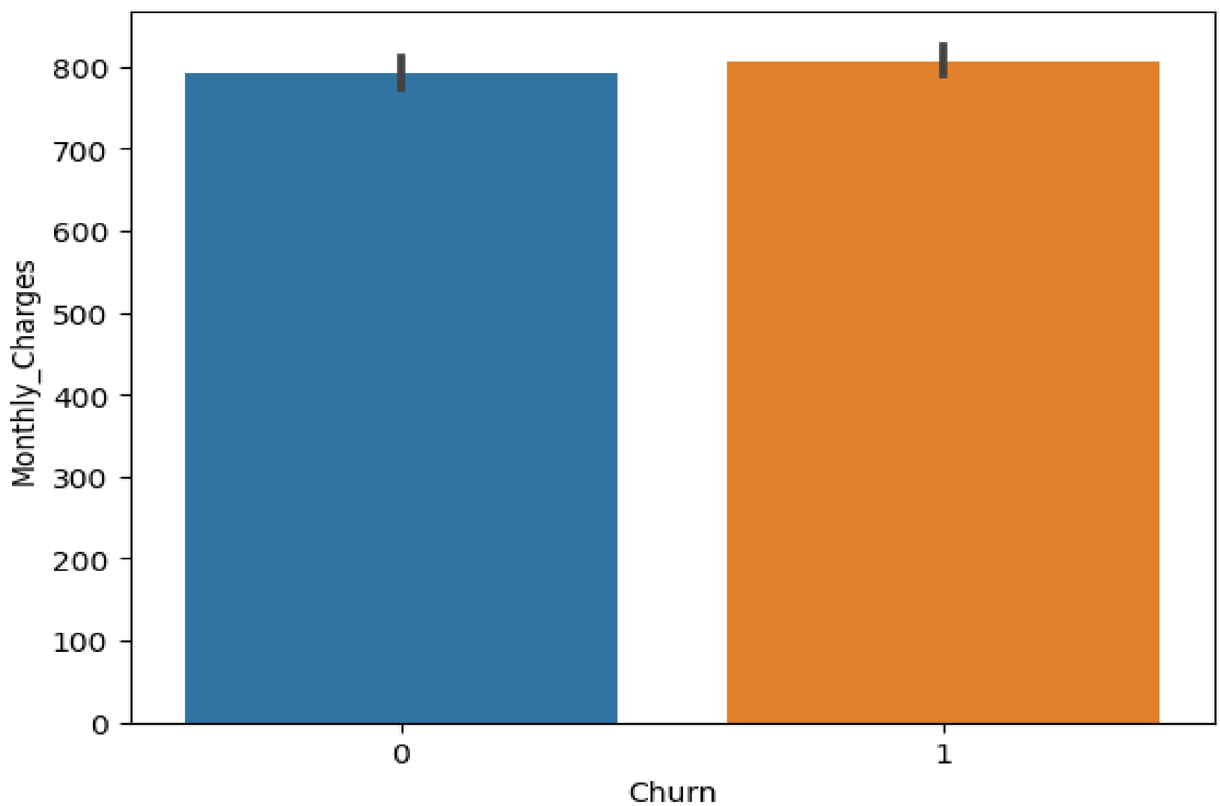


Result 14:

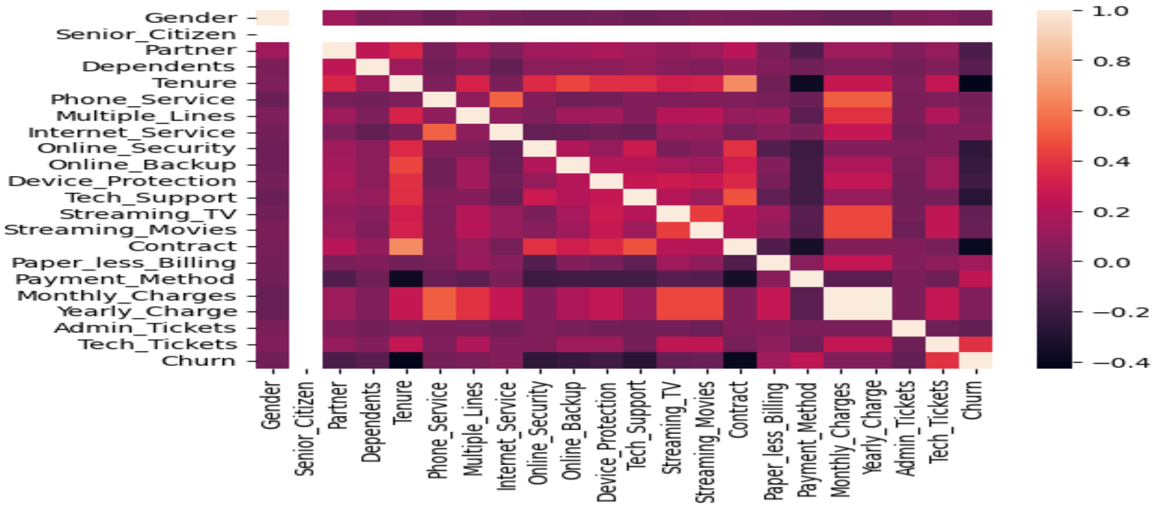




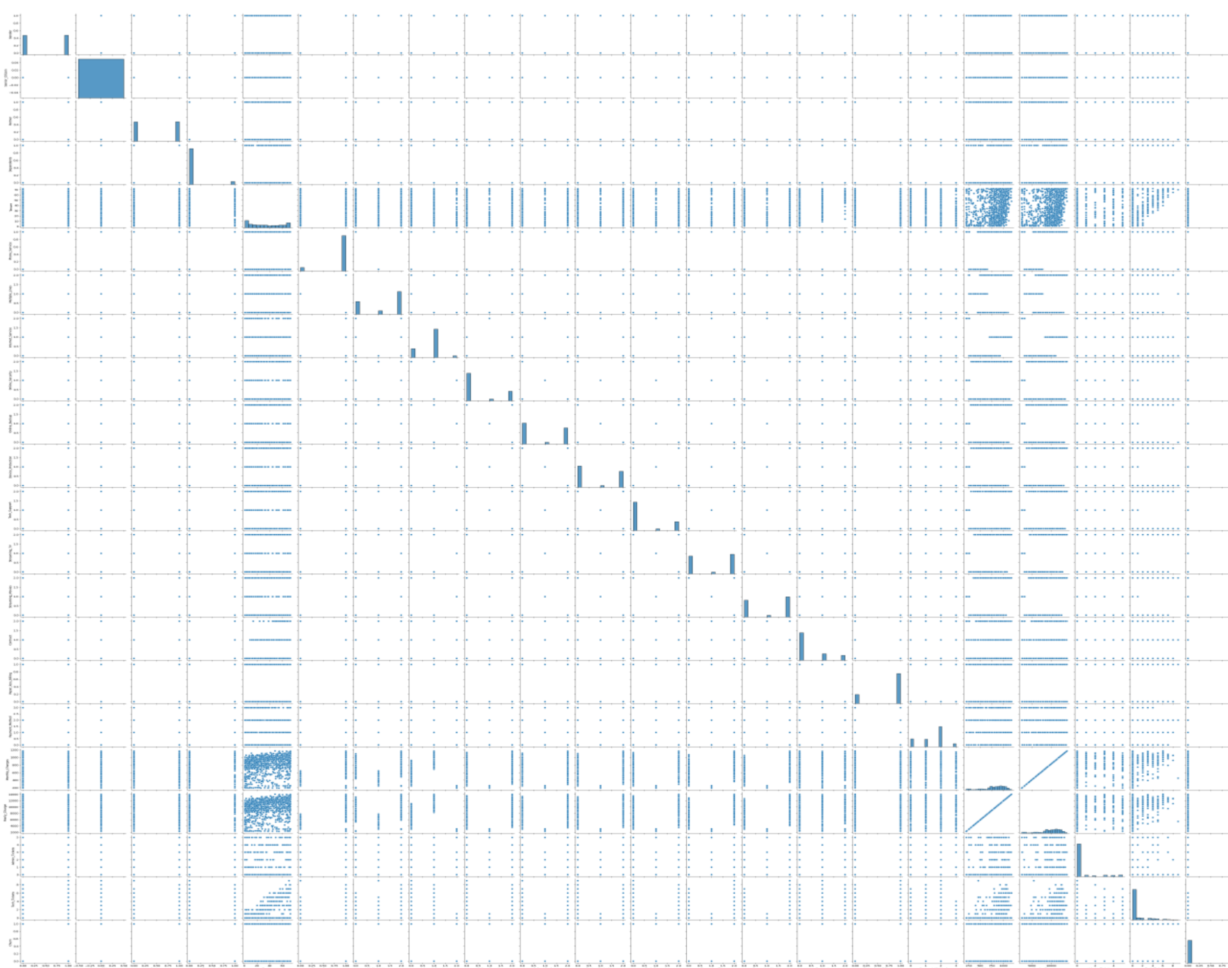
Result 15:



Result 16:



Result 17:



Result 18:

```
array([[1.00000000e+00, 0.00000000e+00, 0.00000000e+00, ...,
        2.00000000e+00, 3.96500000e+02, 4.75800000e+03],
       [0.00000000e+00, 0.00000000e+00, 1.00000000e+00, ...,
        1.00000000e+00, 9.63500000e+02, 1.15620000e+04],
       [1.00000000e+00, 0.00000000e+00, 1.00000000e+00, ...,
        1.00000000e+00, 9.55000000e+02, 1.14600000e+04],
       ...,
       [5.16750610e-01, 0.00000000e+00, 4.83249390e-01, ...,
        2.48324939e+00, 3.75163347e+02, 4.50196016e+03],
       [1.00000000e+00, 0.00000000e+00, 9.58937904e-01, ...,
        1.23186288e-01, 7.10184779e+02, 8.52221735e+03],
       [1.00000000e+00, 0.00000000e+00, 8.00268285e-01, ...,
        1.19973171e+00, 9.01598524e+02, 1.08191823e+04]])
```

Result 19:

```
y_resample
```

```
array([0, 0, 0, ..., 5, 5, 5])
```

```
x.shape, x_resample.shape
```

```
((1142, 19), (5718, 19))
```

```
y.shape, y_resample.shape
```

```
((1142, 1), (5718,))
```

Result 20:

```
([[ 27,  11,  15,  22,  10,  20],
  [ 38, 110,  14,  16,  21,  38],
  [ 38,  16, 159,  44,  29,  18],
  [ 25,   7,  16,  52,  28,  11],
  [ 33,  18,   3,  25,  70,  15],
  [ 35,  19,   7,  32,  33,  69]])
```

Result 21:

```
([[21,  1, 22,  8,  3,  6],
  [39, 94, 44, 25, 31, 39],
  [42, 27, 79, 51, 17, 16],
  [31, 23, 41, 81, 21, 28],
  [27,  8, 22,  5, 78,  1],
  [36, 28,  6, 21, 41, 81]])
```

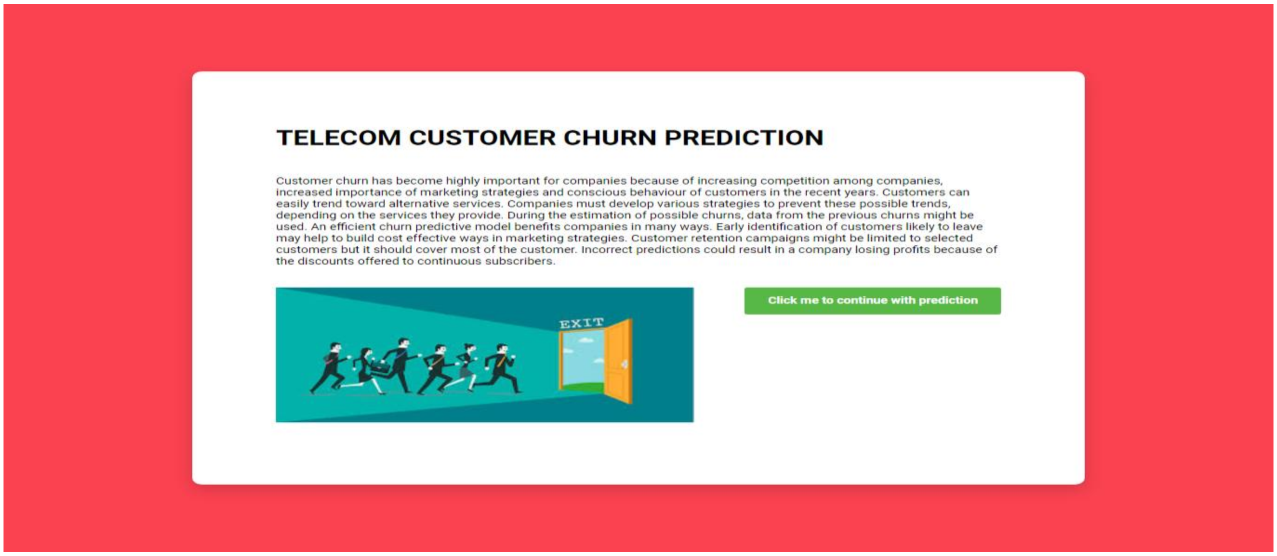
Result 22:

```
([[ 43,   1,  14,   5,   2,   3],
  [ 49, 143,  39,  18,  34,  21],
  [ 16,  19, 107,  32,  17,   3],
  [ 23,   6,  10, 103,  12,  27],
  [ 32,   0,  44,   8, 111,   3],
  [ 33,  12,   0,  25,  15, 114]])
```

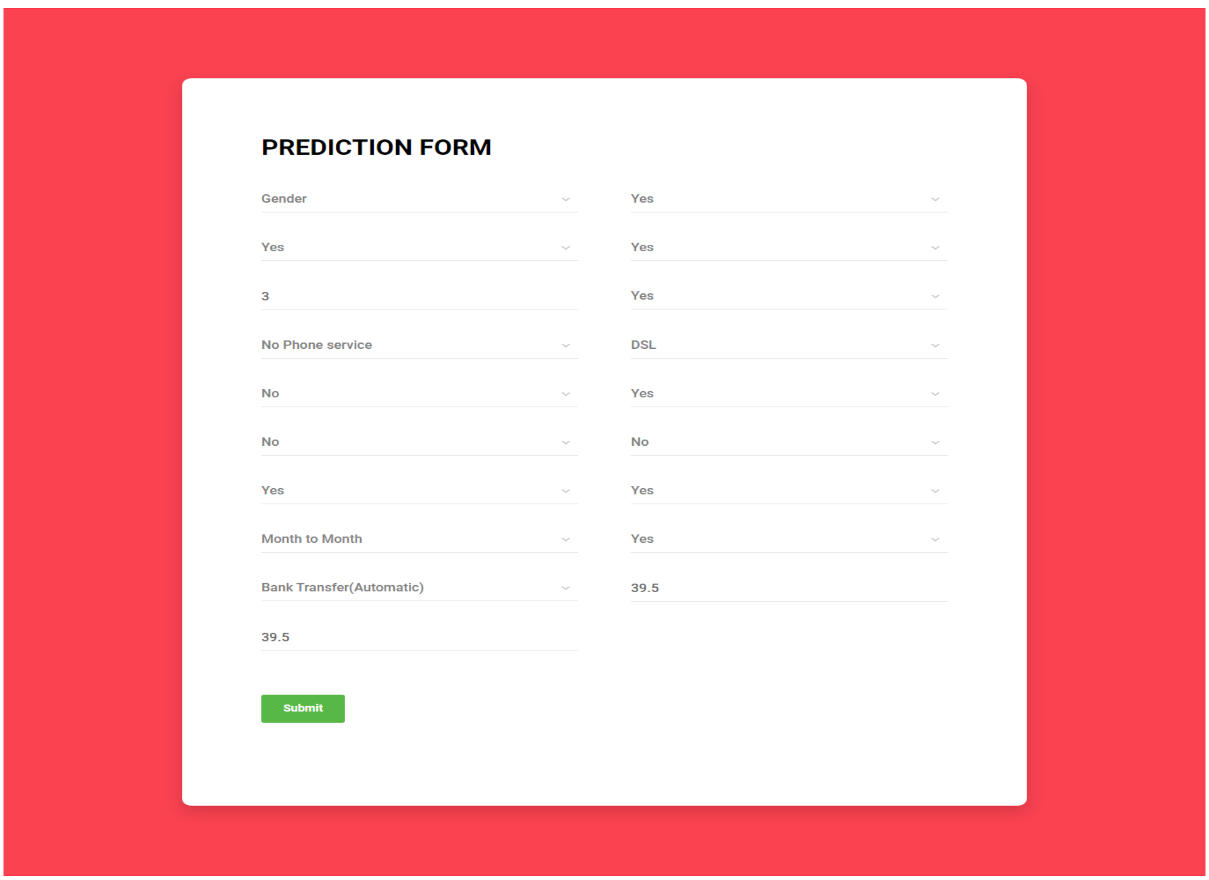
Result 23:

```
[[ 33,   8,   4,  17,  19,  13],
 [ 36, 126,  12,  30,  27,  31],
 [ 34,   8, 185,  29,  32,  20],
 [ 26,   4,   3,  79,  14,  18],
 [ 27,  10,   0,  23,  74,   6],
 [ 40,  25,  10,  13,  25,  83]]
```

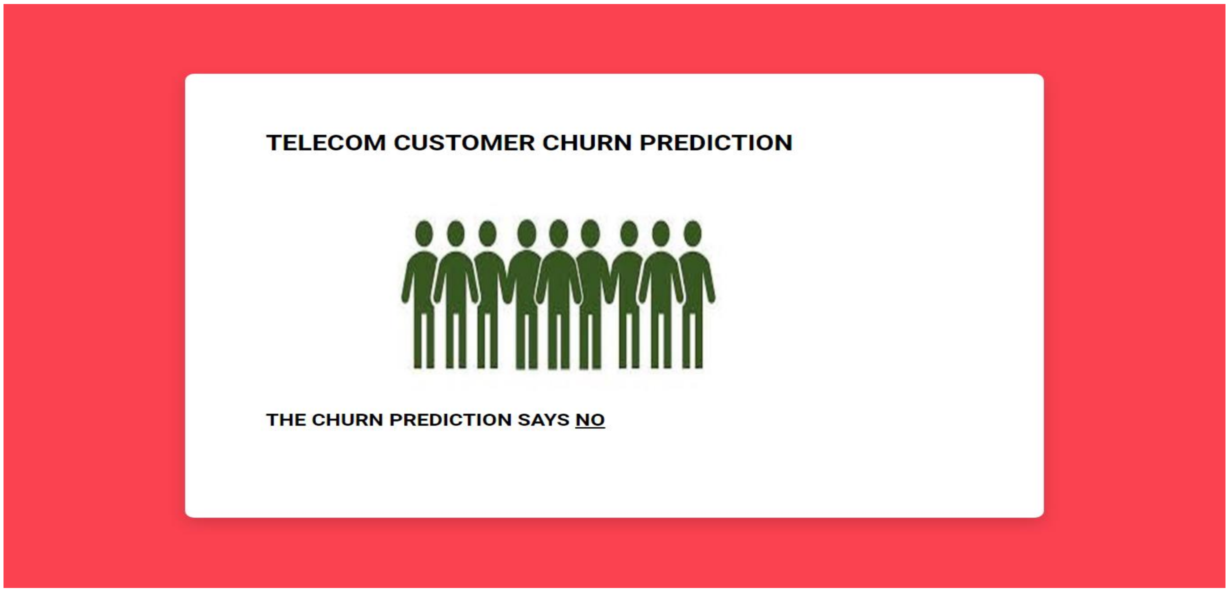
Result 24:


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

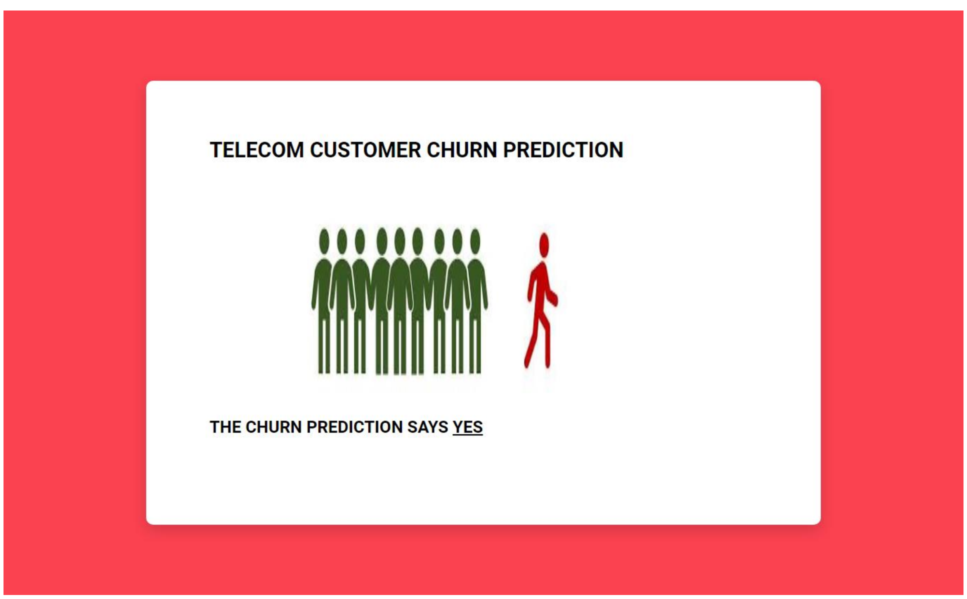
Result 25:


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

Result 26:



Result 27:



CHAPTER-4

4. ADVANTAGES & DISADVANTAGES

ADVANTAGES :

- Improved Accuracy: Machine learning algorithms can analyze large amounts of data and provide highly accurate predictions. By using machine learning, telecom companies can identify customers who are likely to churn with a high degree of accuracy.

- Cost-Effective: Customer acquisition is typically more expensive than customer retention. By identifying at-risk customers and targeting them with retention efforts, telecom companies can reduce churn and save money in the long run.

- Personalization: Machine learning algorithms can analyze customer data and provide insights into individual customer behaviors, preferences, and needs. This allows telecom companies to tailor retention efforts to each customer's specific situation, increasing the chances of success.

- Faster Response: By using machine learning to predict churn, telecom companies can proactively reach out to at-risk customers before they leave. This enables them to address any issues or concerns the customer may have and take steps to retain them.

- Competitive Advantage: Telecom companies that use machine learning for customer retention have a competitive advantage over those that do not. By reducing churn and improving customer loyalty, they can differentiate themselves from competitors and attract new customers.

- Proactive retention: By identifying customers who are at risk of churning before they actually do, telecom companies can be more proactive in their retention efforts. This can include sending personalized retention offers or engaging with the customer to address their concerns and improve their experience with the company.

- Competitive advantage: By leveraging machine learning to improve customer retention, telecom companies can differentiate themselves from their competitors and improve customer loyalty. Customers are more likely to stay with a company that understands their needs and offers personalized solutions to keep them happy.

DISADVANTAGES:

- Data quality: The accuracy of machine learning algorithms depends on the quality of the data they are trained on. If the data is incomplete, inaccurate, or biased, the predictions made by the algorithm may be unreliable. Telecom companies need to ensure that the data they use for training their algorithms is of high quality and representative of their customer base.

- Complexity: Machine learning algorithms can be complex and difficult to understand. This can make it challenging for telecom companies to explain the reasons behind a particular retention strategy to customers who may be skeptical or resistant to the idea of retention offers.

- Ethical concerns: There are ethical concerns around the use of customer data for retention purposes. Telecom companies need to be transparent about the data they are collecting and how it will be used. They also need to ensure that the retention strategies they develop are fair and do not discriminate against certain groups of customers.

- Cost: While customer retention can be less expensive than acquiring new customers, there is still a cost associated with retention efforts. Machine learning algorithms can be expensive to develop and implement, and there may be ongoing costs associated with maintaining and updating them.

- False positives: Machine learning algorithms are not infallible, and there is always a risk of false positives. That is, the algorithm may predict that a customer is at risk of churning when they are not. This can lead to unnecessary retention efforts and may even annoy the customer.

- Privacy concerns: Machine learning algorithms require large amounts of customer data to make accurate predictions. This can raise concerns about privacy and data security, particularly if customers feel that their personal information is being used without their consent or knowledge.

- Bias: Machine learning algorithms are only as unbiased as the data they are trained on. If the data used to train the algorithm is biased or incomplete, it can result in inaccurate predictions and potentially lead to discriminatory treatment of certain customer groups.

CHAPTER-5

5. APPLICATIONS

Intelligent customer retention using machine learning can be used to predict and prevent customer churn in the telecom industry. Here are some potential applications:

- Predicting customer churn: Machine learning algorithms can analyze customer data to predict which customers are at risk of churn. This can be done by looking at factors such as usage patterns, call quality, billing issues, and customer complaints. By identifying customers who are likely to leave, telecom companies can take proactive steps to retain them.

- Identifying key drivers of churn: Machine learning algorithms can also help identify the key drivers of churn. By analyzing data from different sources, such as customer feedback, social media, and customer service interactions, companies can gain insights into the reasons why customers are leaving. This information can be used to improve service quality and customer experience.

- Personalized retention strategies: Machine learning can also help develop personalized retention strategies for individual customers. By analyzing customer data such as past behavior, preferences, and demographics, companies can identify the best retention tactics for each customer. For example, some customers may respond better to discounts or promotions, while others may value personalized customer service or additional features.

- Prioritizing retention efforts: Machine learning algorithms can help prioritize retention efforts by identifying the customers who are most at risk of churn and are most valuable to the company. This can help companies allocate resources more effectively and focus on retaining high-value customers.

- Continuous monitoring: Machine learning can enable continuous monitoring of customer behavior and provide real-time insights into customer sentiment and satisfaction. This can help companies identify potential issues early on and take proactive measures to prevent churn.

Overall, the application of intelligent customer retention using machine learning can help telecom companies improve customer retention and reduce churn, leading to increased revenue and customer satisfaction.

CHAPTER-6

6. CONCLUSION

In conclusion, Intelligent Customer Retention using machine learning is an effective way for telecom companies to enhance their prediction of customer churn and take proactive measures to retain customers. By analyzing customer data, identifying key drivers of churn, developing personalized retention strategies, prioritizing retention efforts, and continuously monitoring customer behavior, telecom companies can improve their customer retention rates and reduce churn. This can lead to increased revenue, improved customer satisfaction, and a competitive advantage in the telecom industry. Therefore, it is essential for telecom companies to leverage the power of machine learning in their customer retention strategies to stay ahead in the market.

Intelligent customer retention using machine learning is a powerful tool that can help telecom companies predict and prevent customer churn. By analyzing customer data, identifying key drivers of churn, developing personalized retention strategies, prioritizing retention efforts, and continuously monitoring customer behavior, companies can improve customer retention and satisfaction while increasing revenue. The use of machine learning in customer retention not only helps companies retain customers but also helps them build strong relationships with customers by providing personalized experiences and better customer service. As such, telecom companies should consider leveraging the power of machine learning to enhance their customer retention efforts and stay ahead of the competition.

CHAPTER-7

7. FUTURE SCOPE

The future scope of intelligent customer retention using machine learning in the telecom industry is vast and promising. Here are some potential areas where this technology can be applied:

- More accurate predictions: As machine learning algorithms become more advanced, they will be able to make even more accurate predictions about customer churn. This could involve the integration of additional data sources, such as social media, to gain more insights into customer behavior.

- Real-time interventions: With the help of machine learning, telecom companies can intervene in real-time to prevent churn. For example, if a customer experiences poor call quality, the company could automatically offer a discount or upgrade to retain the customer.

- Predictive analytics: Machine learning can be used to develop predictive analytics models that identify potential churn risks and opportunities for retention. This could help telecom companies make more informed decisions about where to focus their retention efforts.

- Personalization: Machine learning can be used to provide even more personalized experiences for customers. This could involve the use of chatbots, personalized promotions, and customized product recommendations based on customer behavior.

- Integration with other technologies: The use of machine learning can be integrated with other technologies, such as the Internet of Things (IoT), to gain more insights into customer behavior. For example, data from smart home devices could be used to predict when a customer is most likely to churn.

Overall, the future of intelligent customer retention using machine learning in the telecom industry is exciting. With the potential to make more accurate predictions, intervene in real-time, develop predictive analytics models, provide even more personalized experiences, and integrate with other technologies, companies can improve customer retention and satisfaction while increasing revenue.

CHAPTER-8

8. APPENDIX

8.1 SOURCE CODE

```
import numpy as np

import pandas as pd

import pickle

import matplotlib.pyplot as plt

%matplotlib inline

import seaborn as sns

import sklearn

from sklearn.preprocessing import LabelEncoder, OneHotEncoder

from sklearn.linear_model import LogisticRegression

from sklearn.tree import DecisionTreeClassifier

from sklearn.ensemble import RandomForestClassifier

from sklearn.neighbors import KNeighborsClassifier

from sklearn.svm import SVC

from sklearn.model_selection import RandomizedSearchCV

import imblearn
```

```
from imblearn.over_sampling import SMOTE
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix, f1_score

df = pd.read_excel("/content/Telecom Churn Rate Dataset.xlsx")
df

df['Churn'].value_counts()

df.describe()

df.info()

df.isnull().any()

df.isnull().sum()

df.corr()

sns.heatmap(df.corr(),annot=True)

sns.pairplot(data=df,markers=['^','v'],palette="inferno"

df.head(2)

from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
df["Gender"] = le.fit_transform(df["Gender"])
```

```python
df["Senior_Citizen"] = le.fit_transform(df["Senior_Citizen"])

df["Partner"] = le.fit_transform(df["Partner"])

df["Dependents"] = le.fit_transform(df["Dependents"])

df["Phone_Service"] = le.fit_transform(df["Phone_Service"])

df["Multiple_Lines"] = le.fit_transform(df["Multiple_Lines"])

df["Internet_Service"] = le.fit_transform(df["Internet_Service"])

df["Online_Security"] = le.fit_transform(df["Online_Security"])

df["Online_Backup"] = le.fit_transform(df["Online_Backup"])

df["Device_Protection"] = le.fit_transform(df["Device_Protection"])

df["Tech_Support"] = le.fit_transform(df["Tech_Support"])

df["Streaming_TV"] = le.fit_transform(df["Streaming_TV"])

df["Streaming_Movies"] = le.fit_transform(df["Streaming_Movies"])

df["Contract"] = le.fit_transform(df["Contract"])

df["Paper_less_Billing"] = le.fit_transform(df["Paper_less_Billing"])

df["Payment_Method"] = le.fit_transform(df["Payment_Method"])

df["Churn"] = le.fit_transform(df["Churn"])



for i in df:
  print(df[i].unique())


for i  in df:
  print(df[i].unique())


df.info()

df.corr()

plt.figure(figsize=(12,5))

plt.subplot(1,2,1)

sns.distplot(df["Tenure"])
```

```
plt.subplot(1,2,2)
sns.distplot(df["Monthly_Charges"])

plt.figure(figsize=(12,5))
plt.subplot(1,2,1)
sns.countplot(df["Gender"])
plt.subplot(1,2,2)
sns.countplot(df["Dependents"])

sns.barplot(x="Churn",y="Monthly_Charges",data=df)

sns.heatmap(df.corr(),annot=False)
sns.pairplot(data=df, markers=["^","v"], palette = "inferno")

x= df.iloc[:,0:19].values
y= df.iloc[:,19:20].values

pip install imblearn

from imblearn.over_sampling import SMOTE
smt = SMOTE()
x_resample, y_resample = smt.fit_resample(x,y)
x_resample
y_resample
x.shape, x_resample.shape
y.shape, y_resample.shape

from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(x_resample,y_resample,test_size = 0.2, random_state = 0)
```

```
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x_train = sc.fit_transform(x_train)
x_test = sc.fit_transform(x_test)
x_train.shape

from sklearn.linear_model import LogisticRegression
lr = LogisticRegression(random_state=0)
lr.fit(x_train,y_train)

lr_pred = lr.predict(x_test)

lr_pred

y_test

from sklearn.metrics import accuracy_score
lr_acc = accuracy_score(lr_pred,y_test)

lr_acc

from sklearn.metrics import confusion_matrix
lr_cm = confusion_matrix(lr_pred,y_test)

lr_cm


from sklearn.tree import DecisionTreeClassifier
```

```python
dtc = DecisionTreeClassifier(random_state = 0,criterion= "entropy")
dtc.fit(x_train,y_train)

dtc_pred = dtc.predict(x_test)

dtc_pred

dtc_acc = accuracy_score(dtc_pred,y_test)

dtc_acc

dtc_cm = confusion_matrix(dtc_pred,y_test)
dtc_cm

from sklearn.ensemble import RandomForestClassifier
rfc = RandomForestClassifier(n_estimators= 10,criterion="entropy",random_state=0)
rfc.fit(x_train,y_train)

rfc_pred = rfc.predict(x_test)
rfc_pred

rfc_acc = accuracy_score(rfc_pred,y_test)
rfc_acc
rfc_cm = confusion_matrix(rfc_pred,y_test)
rfc_cm

from sklearn.svm import SVC
svm = SVC(kernel= "linear")
svm.fit(x_train,y_train)
```

```
svm_pred = svm.predict(x_test)

svm_pred

svm_acc = accuracy_score(svm_pred,y_test)

svm_acc


svm_cm = confusion_matrix(svm_pred,y_test)

svm_cm


#svm_pred_own = 
svm.predict(sc.transform([[0,0,1,1,0,0,0,0,1,0,0,1,0,0,1,0,0,1,0,0,1,0,0,1,0,0,1,0,1,0,0,1,1,0,0,456,1,0,3245,456
7]]))

#svm_pred_own


import pickle

pickle.dump(rf,open('churn_rf.pkl','wb'))
```