```
import pandas as pd
import seaborn as sns


dt = pd.read_csv('/content/WA_Fn-UseC_-Telco-Customer-Churn.csv')


dt.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
```

```
for i in dt.columns:
  if dt[i].dtype == 'object':
    dt[i] = dt[i].astype('category').cat.codes


dt.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   int16
 1   gender            7043 non-null   int8
 2   SeniorCitizen     7043 non-null   int64
 3   Partner           7043 non-null   int8
 4   Dependents        7043 non-null   int8
 5   tenure            7043 non-null   int64
 6   PhoneService      7043 non-null   int8
 7   MultipleLines     7043 non-null   int8
 8   InternetService   7043 non-null   int8
```

```
 9   OnlineSecurity    7043 non-null   int8
 10  OnlineBackup      7043 non-null   int8
 11  DeviceProtection  7043 non-null   int8
 12  TechSupport       7043 non-null   int8
 13  StreamingTV       7043 non-null   int8
 14  StreamingMovies   7043 non-null   int8
 15  Contract          7043 non-null   int8
 16  PaperlessBilling  7043 non-null   int8
 17  PaymentMethod     7043 non-null   int8
 18  MonthlyCharges    7043 non-null   float64
 19  TotalCharges      7043 non-null   int16
 20  Churn             7043 non-null   int8
dtypes: float64(1), int16(2), int64(2), int8(16)
memory usage: 302.8 KB
```

```
dt.drop(['customerID'], axis=1, inplace=True)
```

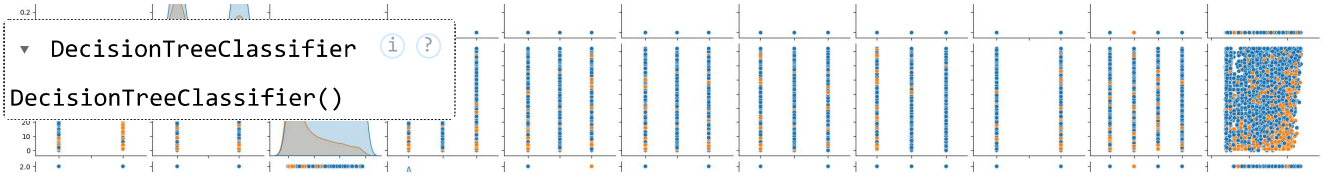
```
dt.corr()
```

```
dt = dt.drop(["gender","Dependents","PhoneService","MultipleLines","InternetService"],axis=1)


dt = dt.drop(["StreamingTV","StreamingMovies","TotalCharges"],axis = 1)


x = dt.drop(["Churn"], axis = 1)
y = dt["Churn"]


from sklearn.model_selection import train_test_split
xtrain, xtest, ytrain, ytest = train_test_split(x, y, test_size = 0.2)


sns.pairplot(data = dt, hue = 'Churn')
```

```
<seaborn.axisgrid.PairGrid at 0x7f803563b850>
```

```
from sklearn.tree import DecisionTreeClassifier
d = DecisionTreeClassifier()
```

```
d.fit(xtrain, ytrain)
```

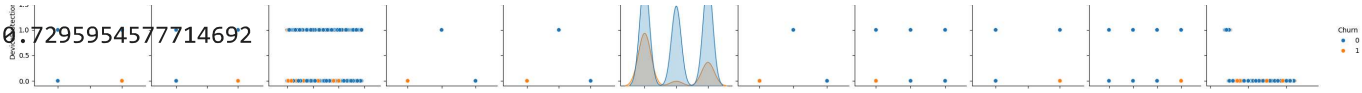



```
ypred = d.predict(xtest)
```

```
from sklearn.metrics import accuracy_score
```

```
accuracy_score(ytest, ypred)
```

```
0.7295954577714692
```




```
import pickle
```

```
pickle.dump(d,open('model.pxl','wb'))
```