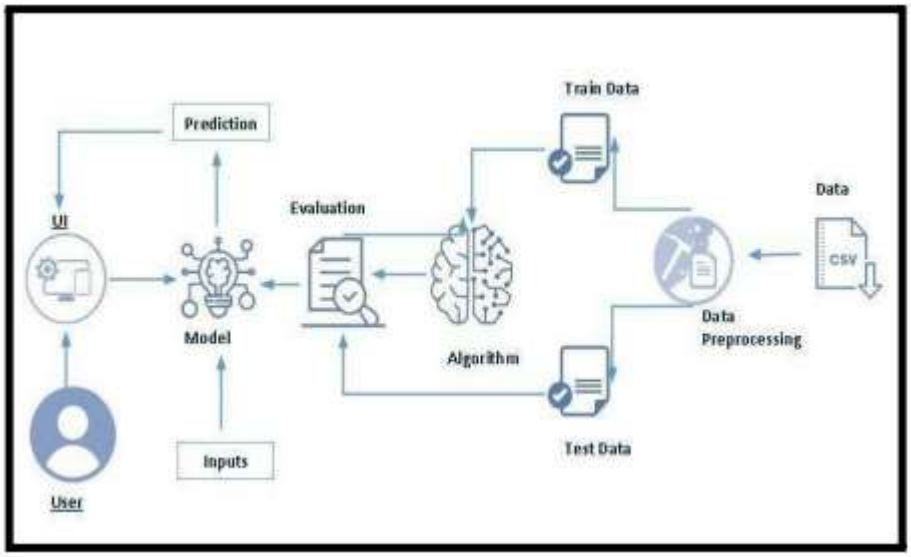
Intelligent Customer Retention: Using Machine Learning for Enhanced Prediction of Telecom Customer Churn

1.1OVERVIEW:

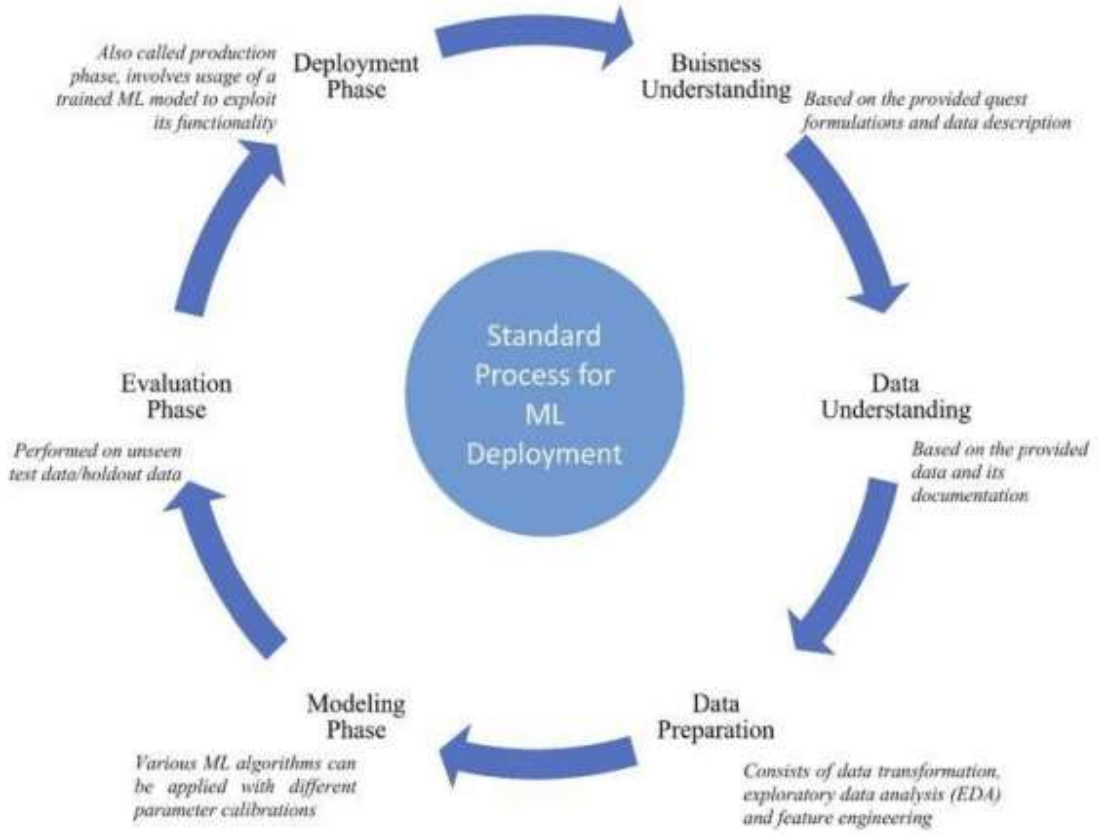
Artificial intelligence (AI) has gradually become accepted by colleges and universities as an effective tool for automating a number of tasks effectively and efficiently.AI-generated emails can remind students about important deadlines, prompt them to register for classes, turn in assignments and pay their fees on time and, in a particularly controversial use. AI-based software is increasingly able to detect plagiarized assignments. AI expands into these core university practices, new concerns are also being raised about the tools threats to personal privacy and its vulnerability to systematicbias.

Technical Architecture:



Technical about the Project:

Machine learning is merely based on predictions made based on experience. It enables machines to make data-driven decisions, which is more efficient than explicitly programming to carry out certain tasks. These algorithms are designed in a fashion that gives exposure to new data that can help organizations learn and improve their strategies.



A PROJECT DESCRIPTION:

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.

Project Flow:

- User interacts with the UI to enter the input.
- Entered input is analysed by the model which is integrated.
- Once model analyses the input the prediction is showcased on the UI To

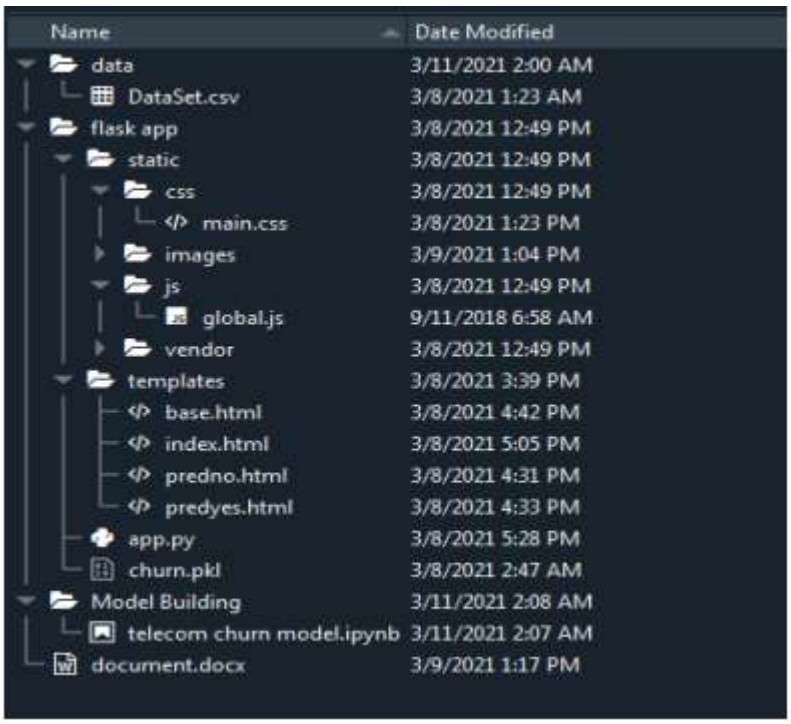
accomplish this, we have to complete all the activities listed below,

- Define Problem / Problem Understanding
  - Specify the business problem
  - Business requirements ○ Literature Survey
  - Social or Business Impact.
- Data Collection & Preparation
  - Collect the dataset
  - Data Preparation
- Exploratory Data Analysis
  - Descriptive statistical
  - Visual Analysis
- Model Building
  - Training the model in multiple algorithms
  - Testing the model
- Performance Testing & Hyperparameter Tuning
  - Testing model with multiple evaluation metrics
  - Comparing model accuracy before & after applying hyperparameter tuning

- Model Deployment
    - Save the best model
    - Integrate with Web Framework
- Project Demonstration & Documentation
    - Record explanation Video for project end to end solution
    - Project Documentation-Step by step project development procedure

Project Structure:

Create a Project folder which contains files as shown below



- A python file called app.py for server side scipting.
- We need the model which is saved and the saved model in this content is churn.pkl
- Templates folder which contains base.HTML file, index.HTML file, predyes.HTML , predno.HTML file.
- Static folder which contains css folder which contains main.css , js folder which contains global.js , images folder and vendor folder.

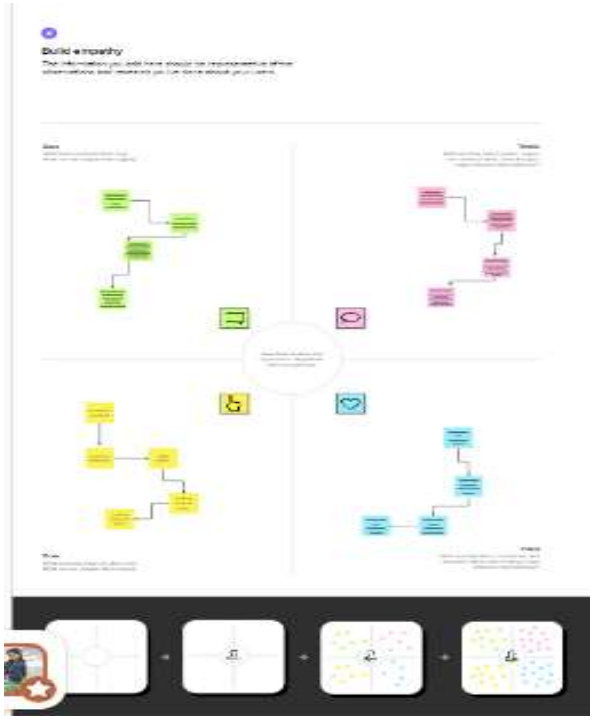
1.2 PURPOSE:

Machine Learning has an additional benefit of processing large chunks of data that is sometimes tire some for men to do and eventually lead to a failure in making the right decision .It is easily adaptable to new and complex data. After processing the data, it is capable of analyzing any flaw so r errors. These also help in creating effective plans of Actions for improvement. There is acceleration between inputs and outputs in the process of decision-making. These points are extremely useful for ventures that work mainly around risk management.

2. PROBLEM DEFINITION AND DESIGN THINKING

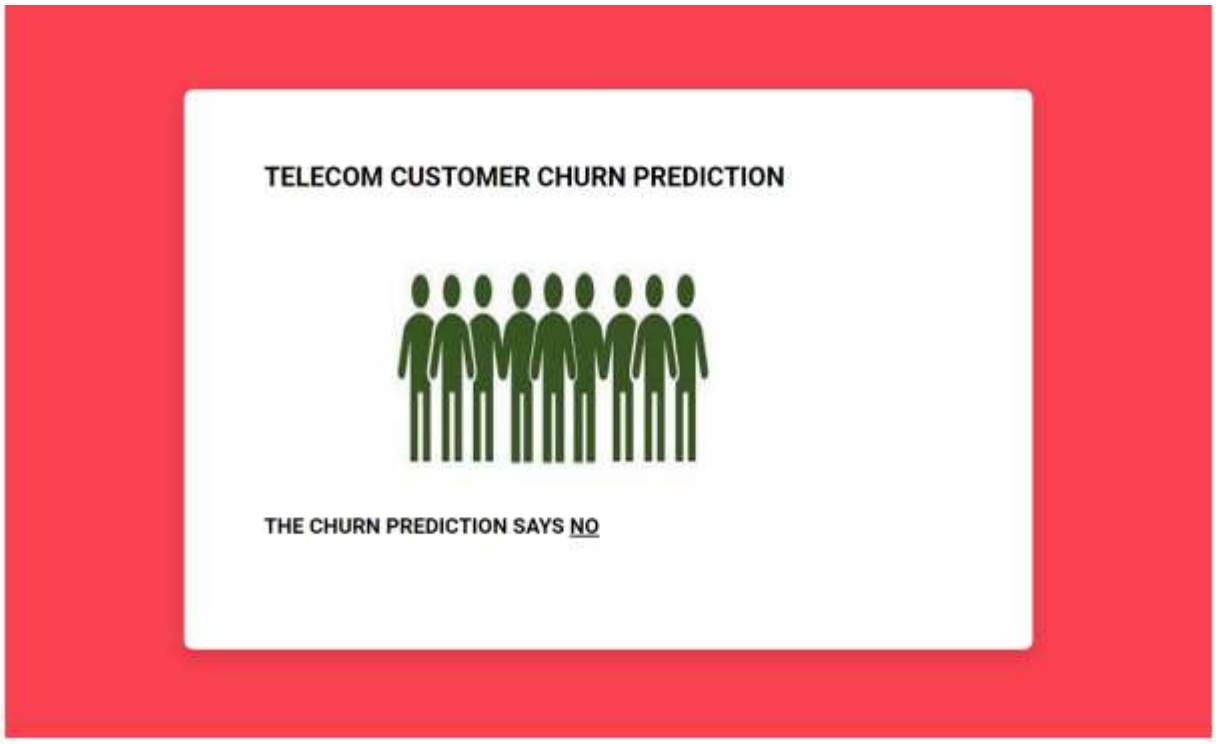
Machine learning has become an increasingly popular too recent years, given its ability to automatically detect pattern sin data and make predictions about future events. This can be extremely useful for making decisions in a wide range of domains, from financial trading to medical diagnoses.

2.1 EMPATHY MAP



2.2 Ideation and Brainstorming Map:

3. RESULT:



4. ADVANTAGE AND DISADVANTAGE:

Advantage:

1. Providing better information: Since machine learning technology can sift through extremely large amounts of data, it is able to also provide better information to decision makers.

2. Automating the process: In many industries, it is simply not possible for human beings to make optimal decisions all of the time .This is especially true in industries where the data is constantly changing, such as financial markets .In these cases, machine learning algorithm can be used to automatically make decisions as trends change and evolve.

3. Improving accuracy: By identifying patterns in data that humans may not be ableto see, machine learning can drastically improve the accuracy of its predictions. It can also create models that simulate different decision scenarios and help identify the best course of action. And as new data becomes available, machine learning can be used to constantly update and refine decision models.

Disadvantage:

1. Data Acquisition

Machine Learning requires massive data sets to train on , and these should be inclusive/unbiased, and of good quality .There can also be times where they must wait for new data to be generated.

2. Time and Resources

ML needs enough time to let the algorithms learn and develop enough to fulfill their purpose with a considerable amount of accuracy and relevancy .It also needs massive resources to function.This can mean additional requirements of computer power for you.

3. Interpretation of Results

Another major challenge is the ability to accurately interpret results generated by the algorithms .You must also carefully choose the algorithms for your purpose.

4. High error-susceptibility

Machine Learning is autonomous but highly susceptible to errors .Suppose you train an algorithm with data sets small enough on it be inclusive .You end up with biased predictions coming from a biased training set . This leads to irrelevant advertisements being displayed to customers .In the case of ML, such blunder scan set off a chain of errors that can go undetected for long periods of time . And when they do get noticed , it takes quite some time to recognize the source of the issue ,and even longer to correct it.

5.APPLICATION

1. Decisions in business operations

Machine Leaíning algorithms come to the íescue in aíeas built on a constant flow of heterogeneous data, whether it is several financial reports, payrolls, procurement, the analysis of employee productivity, or predicting further churn rates. Overall, AI, in terms of inner business processes, is able to leverage business intelligence and make a company data-driven in many aspects, including decision making.

2. Complex problem-solving

The potential of AI in decision making is robust, but you can solve multilayer and complex problems, too. Artificial Intelligence here gathers tons of different data and conducts an inter disciplinary study. Eventually, there's a way to leverage anything from product development stages to digital marketing approaches of product píomotion. Also, it's a way to optimize various types of predictions and risk management. For example, you can predict and optimize píicing with the help of AI tools.

3. Strategic changes

AI allows better planning of production, managing all restrictions, reducing shortcomings in operations, and improving manufacturing. It also helps to anticipate and adequately plan product customization, enhance postponement processes, and maintain efficiency with high levels of customer satisfaction

4. Customer-related decisions

Al can be valuable for customer service management, personalized customer communication, evaluation of customer behavior, predicting consumer trends and patterns. Artificial intelligence enables automatic recognition and profiling of potential customers.

5. Performance assessment

Firstly, it relates to people's performance evaluation and afterward decisions. AI is capable of minimizing human errors and making employee performance data moíe transparent.AI can also recommend online courses, training, and development programs to employees based on their performance history.

6.CONCLUSION

Customer retention was determined as the way to keep the customers in house, maintain the customer relationship and the improvement of the level of customer service. The improvement of the level of customer service was seen as the end result of the activities the customer dissatisfaction causes. Overall, being part of the customer retention was experienced to be hard. Coping with the anger customer might have when being in contact with the company was experienced either harder to manage or easier to manage depending on the root cause.

Listening to them and understanding their needs. Never miss out on any opportunity to surprise your customers. Try to deliver better service at a lower cost, do more than they were expecting. Treat them as special people and you will for sure improve customer retention programs.

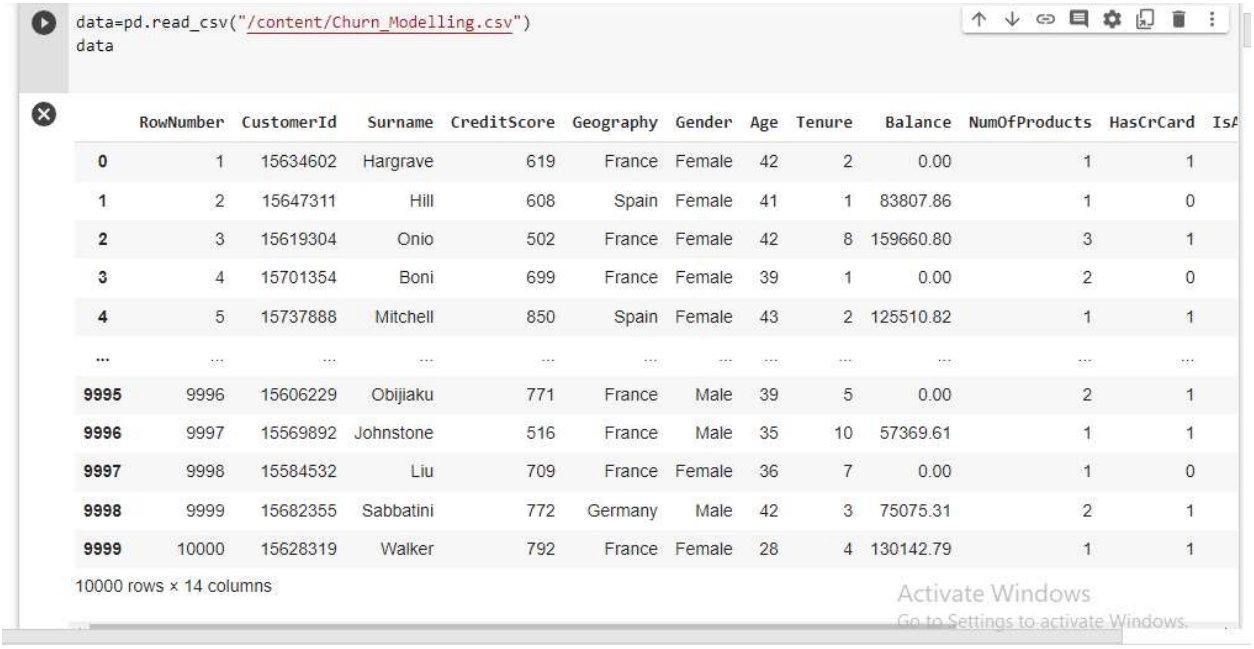
7. FUTURE SCOPE

With the help of machine learning services like SDKs and APIs, developers are able to include and hone the intelligent capabilities into their applications. This will empower machines to apply the various things they come across, and accordingly carry out an array of duties like vision recognition, speech detection, and understanding of speech and dialect.

8. APPENDIX

SourceCode:

```
import pandas as pd
import numpy as np
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_s
core
```



```
data=pd.read_csv("/content/Churn_Modelling.csv")
data
```

	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsA
0	1	15634602	Hargrave	619	France	Female	42	2	0.00	1	1	
1	2	15647311	Hill	608	Spain	Female	41	1	83807.86	1	0	
2	3	15619304	Onio	502	France	Female	42	8	159660.80	3	1	
3	4	15701354	Boni	699	France	Female	39	1	0.00	2	0	
4	5	15737888	Mitchell	850	Spain	Female	43	2	125510.82	1	1	
...	...	...	...	...	...	...	...	...	...	...	...	...
9995	9996	15606229	Obijiaku	771	France	Male	39	5	0.00	2	1	
9996	9997	15569892	Johnstone	516	France	Male	35	10	57369.61	1	1	
9997	9998	15584532	Liu	709	France	Female	36	7	0.00	1	0	
9998	9999	15682355	Sabbatini	772	Germany	Male	42	3	75075.31	2	1	
9999	10000	15628319	Walker	792	France	Female	28	4	130142.79	1	1	

10000 rows × 14 columns

data=data.rename(columns={'Chance of Admit':`Chance of Admit'})

```
data.isnull().sum()
```

```
RowNumber              0
CustomerId             0
Surname                0
CreditScore            0
Geography          10000
Gender                 0
Age                    0
Tenure                 0
Balance                0
NumOfProducts          0
HasCrCard              0
IsActiveMember         0
EstimatedSalary        0
Exited                 0
dtype: int64
```

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   RowNumber        10000 non-null  int64  
 1   CustomerId       10000 non-null  int64  
 2   Surname          10000 non-null  object 
 3   CreditScore      10000 non-null  int64  
 4   Geography        10000 non-null  object 
 5   Gender           10000 non-null  object 
 6   Age              10000 non-null  int64  
 7   Tenure           10000 non-null  int64  
 8   Balance          10000 non-null  float64
 9   NumOfProducts    10000 non-null  int64  
 10  HasCrCard        10000 non-null  int64  
 11  IsActiveMember   10000 non-null  int64  
 12  EstimatedSalary  10000 non-null  float64
 13  Exited           10000 non-null  int64  
dtypes: float64(2), int64(9), object(3)
memory usage: 1.1+ MB
```

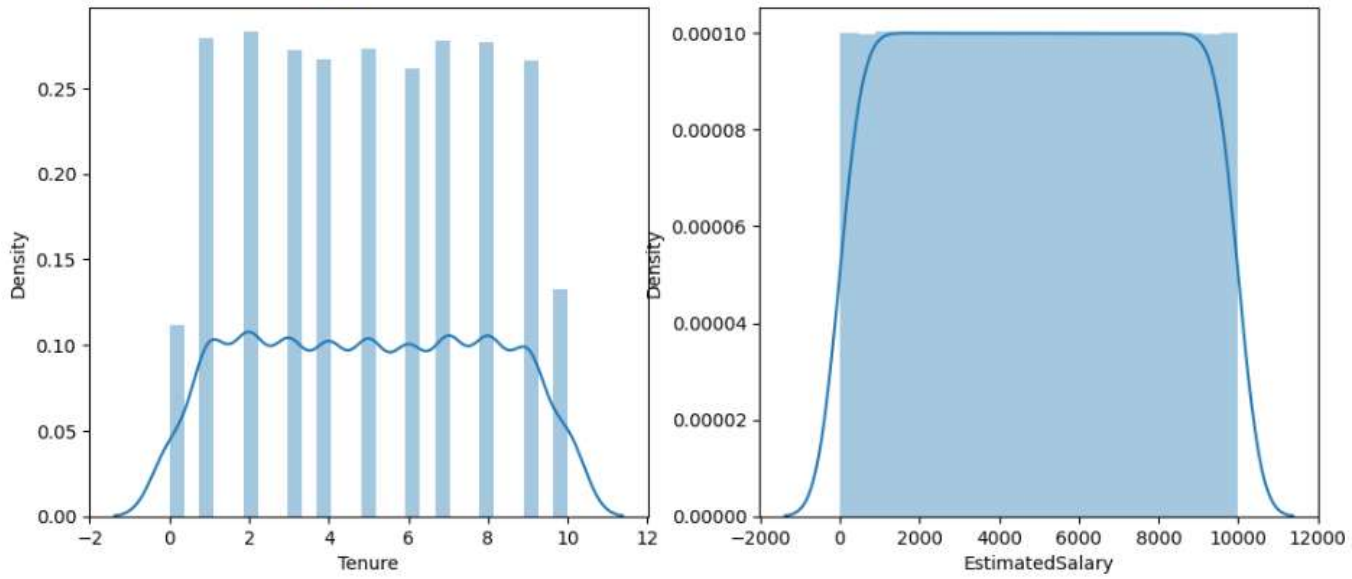
```
data.describe()
```

	RowNumber	CustomerId	Surname	CreditScore	Gender	Age	Tenure	Balance	NumOfProd
count	10000.00000	10000.00000	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000	10000.00
mean	4999.50000	4999.50000	1507.774200	259.584600	0.545700	20.920600	5.012800	2036.788100	0.53
std	2886.89568	2886.89568	846.204311	96.496107	0.497932	10.482065	2.892174	2125.232536	0.58
min	0.00000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00
25%	2499.75000	2499.75000	773.750000	193.000000	0.000000	14.000000	3.000000	0.000000	0.00
50%	4999.50000	4999.50000	1542.000000	261.000000	1.000000	19.000000	5.000000	1383.500000	0.00
75%	7499.25000	7499.25000	2238.250000	327.000000	1.000000	26.000000	7.000000	3882.250000	1.00
max	9999.00000	9999.00000	2931.000000	459.000000	1.000000	69.000000	10.000000	6381.000000	3.00

Visual analysis

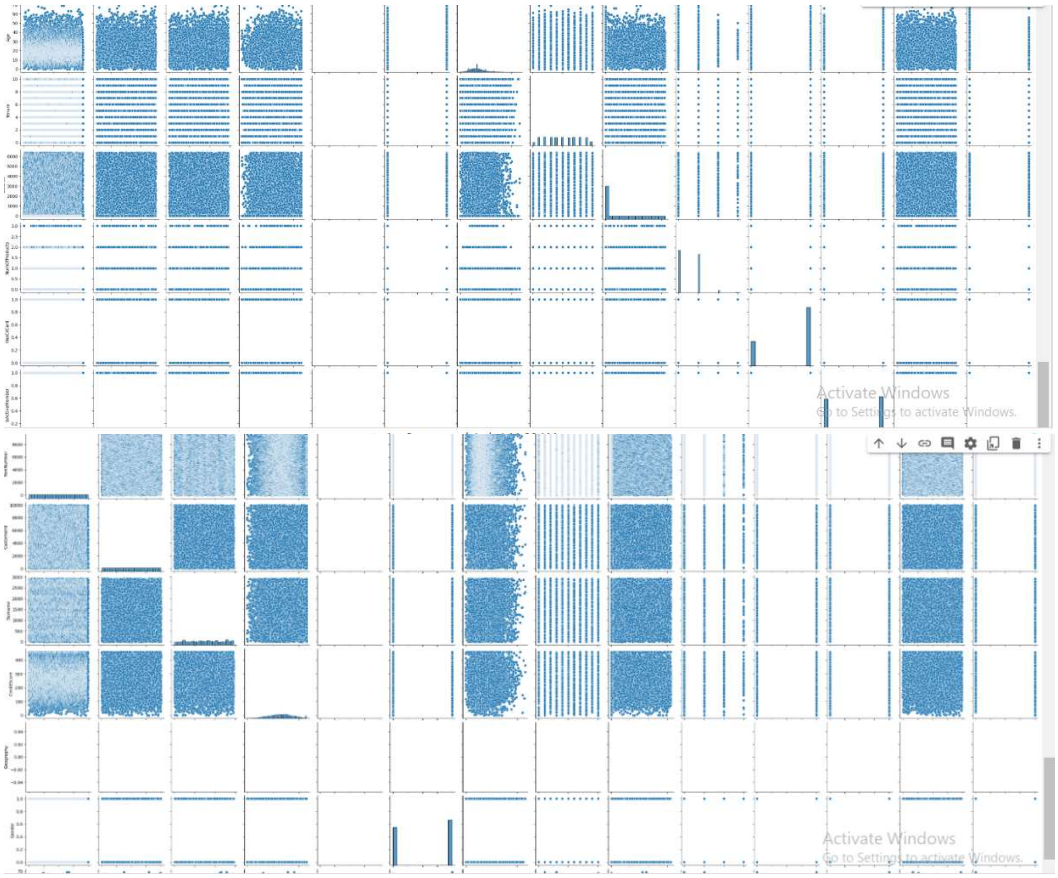
Univariate

```
sns.distplot(data["EstimatedSalary"])
<Axes: xlabel='EstimatedSalary', ylabel='Density'>
```



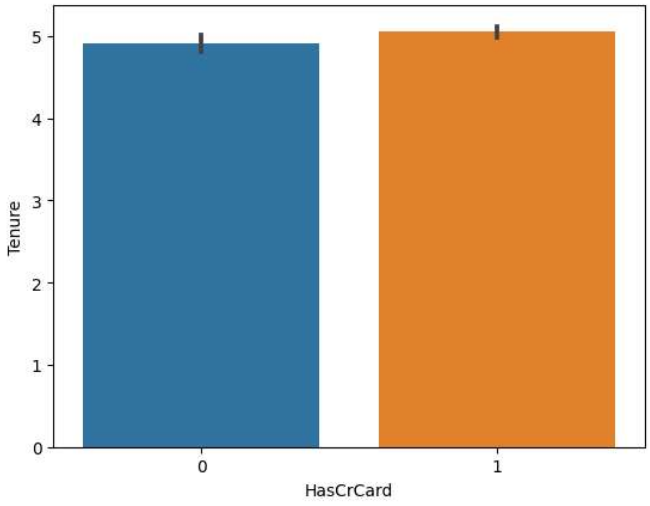
Analysis

Bivariate analysis



```
sns.barplot(x="HasCrCard",y="Tenure",data=data)
```

```
<Axes: xlabel='HasCrCard', ylabel='Tenure'>
```



```
from sklearn.linear_model.logistic import LogisticRegression
cls =LogisticRegression(random_state =0)

lr=cls.fit(x_train, y_train)
```

```
C:\Users\Tulasi\anaconda3\lib\site-packages\sklearn\utils\validation.py:760: DataConversionWarn
array was expected. Please change the shape of y to (n_samples, ), for example using ravel().
  y = column_or_1d(y, warn=True)
```

```
y_pred =lr.predict(x_test)
y_pred
```

ModelDeployment::Savethebestmodel:

```
# Save the model in HDF5 format
  model.save('model.h5')
```

```
  File "/tmp/ipykernel_27/3721887466.py", line 2
    model.save('model.h5')
    ^
IndentationError: unexpected indent
```

IntegratewithWebFramework:

```
import numby as np
from flask import Flask, request,jsonify, render_template
import pickle
app = Flask(__name__)
# Import necessary libraries
from tensorflow.keras.models import load_model

#model = pickle.load(open('university.pkl', 'rb'))
```

```
---------------------------------------------------------------------------
ModuleNotFoundError                       Traceback (most recent call last)
/tmp/ipykernel_27/4117061979.py in <module>
----> 1 import numby as np
      2 from flask import Flask, request,jsonify, render_template
      3 import pickle
      4 app = Flask(__name__)
      5 # Import necessary libraries

ModuleNotFoundError: No module named 'numby'
```

```
#load model trained model
#Load your trained model
model = load_ model('model.h5')
```

```
  File "/tmp/ipykernel_27/552793924.py", line 3
    model = load_ model('model.h5')
                        ^
SyntaxError: invalid syntax
```

```
@app.route('/')
def home():
  return render_template('Demo2.html')
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
/tmp/ipykernel_27/3305153883.py in <module>
----> 1 @app.route('/')
      2 def home():
      3   return render_template('Demo2.html')

NameError: name 'app' is not defined
```

RetrievesthevaluefromUI:

```
@app.route('/')
def home():
  return render_template('Demo2.html')

@app.route('/y_predict',methods=['POST'])
def y_predict():

  For rendering results on HTML GUI

#min max scaling
min1=[290.0, 92.0, 1.0, 1.0, 1.0, 6.0, 0.0]
max1=[340.0, 120.0, 5.0, 5.0, 9.92, 1.0]
k= [float(x) for x in request.form.values()]
p=[]
for i in range(7):
    l=(k[i]-min1[i])/(max1[i]-min1[i])
    p.append(l)
prediction = model.predict([p])
print(prediction)
output=prediction[0]
if(output<=False):
  return render_template('noChance.html', prediction_text='You Dont have a chance of getting
else:
   return render_template('Chance.html', prediction_text='You  have a chance of getting admis
if__name__ == "__main__":
   app.run(debug=False)
```

```
  File "/tmp/ipykernel_27/3019622180.py", line 8
    For rendering results on HTML GUI
            ^
SyntaxError: invalid syntax
```