The Superior University

Name:Fatima Nadeem	Roll No:076	Course: PAI Lab
Semester:4rd	Section:4B	Department:SE
Submitted To:Sir Rasikh	Total Marks: 10	Date:

Lab Task 02

Spaceship Titanic

Table of content.

Introduction:

In this competition your task is to predict whether a passenger was transported to an alternate dimension during the *Spaceship Titanic*'s collision with the spacetime anomaly. To help you make these predictions, you're given a set of personal records recovered from the ship's damaged computer system.

Problem statement:

To help rescue crews and retrieve the lost passengers, you are challenged to predict which passengers were transported by the anomaly using records recovered from the spaceship’s damaged computer system.

Step1:

Read the csv file train and test file.By using pandas library.

```
[51]: import pandas as pd
```

```
[52]: train_data=pd.read_csv('train.csv')
train_data
```

[52]:

	PassengerId	HomePlanet	CryoSleep	Cabin	Destination	Age	VIP	RoomService	FoodCourt	ShoppingMall	Spa	VRDeck	Name	Transpo
0	0001_01	Europa	False	B/0/P	TRAPPIST-1e	39.0	False	0.0	0.0	0.0	0.0	0.0	Maham Ofracculy	F
1	0002_01	Earth	False	F/0/S	TRAPPIST-1e	24.0	False	109.0	9.0	25.0	549.0	44.0	Juanna Vines	
2	0003_01	Europa	False	A/0/S	TRAPPIST-1e	58.0	True	43.0	3576.0	0.0	6715.0	49.0	Altark Susent	F
3	0003_02	Europa	False	A/0/S	TRAPPIST-1e	33.0	False	0.0	1283.0	371.0	3329.0	193.0	Solam Susent	F
4	0004_01	Earth	False	F/1/S	TRAPPIST-1e	16.0	False	303.0	70.0	151.0	565.0	2.0	Willy Santantines	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
8688	9276_01	Europa	False	A/98/P	55 Cancri e	41.0	True	0.0	6819.0	0.0	1643.0	74.0	Gravior Noxnuther	F
8689	9278_01	Earth	True	G/1499/S	PSO J318.5-22	18.0	False	0.0	0.0	0.0	0.0	0.0	Kurta Mondalley	F
8690	9279_01	Earth	False	G/1500/S	TRAPPIST-1e	26.0	False	0.0	0.0	1872.0	1.0	0.0	Fayey Connon	
8691	9280_01	Europa	False	E/608/S	55 Cancri e	32.0	False	0.0	1049.0	0.0	353.0	3235.0	Celeon Hontichre	F
8692	9280_02	Europa	False	E/608/S	TRAPPIST-1e	44.0	False	126.0	4688.0	0.0	0.0	12.0	Propsh Hontichre	

```
: test_data=pd.read_csv('test.csv')
test_data.head()
```

:

	PassengerId	HomePlanet	CryoSleep	Cabin	Destination	Age	VIP	RoomService	FoodCourt	ShoppingMall	Spa	VRDeck	Name
0	0013_01	Earth	True	G/3/S	TRAPPIST-1e	27.0	False	0.0	0.0	0.0	0.0	0.0	Nelly Carsoning
1	0018_01	Earth	False	F/4/S	TRAPPIST-1e	19.0	False	0.0	9.0	0.0	2823.0	0.0	Lerome Peckers
2	0019_01	Europa	True	C/0/S	55 Cancri e	31.0	False	0.0	0.0	0.0	0.0	0.0	Sabih Unhearfus
3	0021_01	Europa	False	C/1/S	TRAPPIST-1e	38.0	False	0.0	6652.0	0.0	181.0	585.0	Meratz Caltilter
4	0023_01	Earth	False	F/5/S	TRAPPIST-1e	20.0	False	10.0	0.0	635.0	0.0	0.0	Brence Harperez

Step2:

EDA Steps ,Preprocessing and Feature engineering.

55]:
```
train_data.info()
train_data.isnull
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8693 entries, 0 to 8692
Data columns (total 14 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   PassengerId   8693 non-null   object 
 1   HomePlanet    8492 non-null   object 
 2   CryoSleep     8476 non-null   object 
 3   Cabin         8494 non-null   object 
 4   Destination   8511 non-null   object 
 5   Age           8514 non-null   float64
 6   VIP           8490 non-null   object 
 7   RoomService   8512 non-null   float64
 8   FoodCourt     8510 non-null   float64
 9   ShoppingMall  8485 non-null   float64
 10  Spa           8510 non-null   float64
 11  VRDeck        8505 non-null   float64
 12  Name          8493 non-null   object 
 13  Transported   8693 non-null   bool   
dtypes: bool(1), float64(6), object(7)
memory usage: 891.5+ KB
```

55]:
```
<bound method DataFrame.isnull of      PassengerId HomePlanet CryoSleep     Cabin    Destination   Age    VIP  \
0        0001_01     Europa     False     B/0/P    TRAPPIST-1e  39.0  False   
1        0002_01      Earth     False     F/0/S    TRAPPIST-1e  24.0  False   
2        0003_01     Europa     False     A/0/S    TRAPPIST-1e  58.0   True   
3        0003_02     Europa     False     A/0/S    TRAPPIST-1e  33.0  False   
4        0004_01      Earth     False     F/1/S    TRAPPIST-1e  16.0  False   
...          ...        ...       ...       ...            ...   ...    ...   
8688     9276_01     Europa     False    A/98/P    55 Cancri e  41.0   True   
8689     9278_01      Earth      True  G/1499/S  PSO J318.5-22  18.0  False   
```

5]:
```
y = train_data[['PassengerId', 'Transported','Destination']]
```

5]:
```
features = ['PassengerId', 'Transported']
```

7]:
```
import numpy as np


from sklearn.svm import SVC
from sklearn.metrics import accuracy_score
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
```

3]:
```
# def dataEncoder(cols):
#     for i in cols:
#         dataLabelEncoder = LabelEncoder()
#         train_data[i] = dataLabelEncoder.fit_transform(train_data[i])

# columns = ['PassengerId']
# dataEncoder(columns)
```

9]:
```
 train_data['Transported'] = train_data['Transported'].astype('int64')
```

9]:
```
x = train_data[features]
```

```
x = train_data[features]
```

```
def dataEncoder(cols):
    for i in cols:
        dataLabelEncoder = LabelEncoder()
        train_data[i] = dataLabelEncoder.fit_transform(train_data[i])

columns = ['Destination']
dataEncoder(columns)
```

```
y = train_data['Transported'].values
```

```
y = train_data['Transported']
```

```
model_svc = SVC()
model_svc.fit(x, y)

print(model_svc)
```

```
SVC()
```

```
df = pd.read_csv("train.csv")
selected_features = df[['PassengerId', 'Transported']]
selected_features.to_csv("new.csv", index=False)

print("New CSV file created successfully!")
```

```
New CSV file created successfully!
```

```
test11_data=pd.read_csv('new.csv')
```

Step3:

Training the model of svc on two attributes and predicating the transported people and completing the Competition.

```
x = train_data[features]
```

```
def dataEncoder(cols):
    for i in cols:
        dataLabelEncoder = LabelEncoder()
        train_data[i] = dataLabelEncoder.fit_transform(train_data[i])

columns = ['Destination']
dataEncoder(columns)
```

```
y = train_data['Transported'].values
```

```
y = train_data['Transported']
```

```
model_svc = SVC()
model_svc.fit(x, y)

print(model_svc)
```

```
SVC()
```

```
df = pd.read_csv("train.csv")
selected_features = df[['PassengerId', 'Transported']]
selected_features.to_csv("new.csv", index=False)

print("New CSV file created successfully!")
```

```
New CSV file created successfully!
```

```
test11_data=pd.read_csv('new.csv')
```

```
SVC()
```

```
df = pd.read_csv("train.csv")
selected_features = df[['PassengerId', 'Transported']]
selected_features.to_csv("new.csv", index=False)

print("New CSV file created successfully!")
```

```
New CSV file created successfully!
```

```
test11_data=pd.read_csv('new.csv')
```

```

model_predictions = model_svc.predict(test11_data)
print("Predictions on Test Data:", model_predictions)
```

```
Predictions on Test Data: [0 0 0 ... 0 0 0]
```

Submission on kaggle:

Predictions on Test Data: [0 0 0 ... 0 0 0]

[90]:
```
submission10=pd.DataFrame({'PassengerId':test11_data['PassengerId'],'Transported':model_predictions})
submission10.to_csv('submission10.csv',index=False)
print("submission successfully")
```

```
submission successfully
```

[ ]:

Submissions

All | Successful | Errors | Recent

Submission and Description	Public Score
submission10.csv Complete · 1mo ago	**0.53682**
submission0-7.csv Error · 1mo ago	
submission0-6.csv Complete · 1mo ago	**0.00000**