Welcome To Colaboratory - Colal | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help

Comment Share

+ Code + Text

RAM Disk

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier,GradientBoostingClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import f1_score
from sklearn.metrics import classification_report,confusion_matrix
import warnings
import pickle
from scipy import stats
warnings.filterwarnings('ignore')
plt.style.use('fivethirtyeight')
```

0s

```
[2] data=pd.read_csv("/content/Data_Train_1.csv")
    data.head()
```

1s

	Airline	Date_of_Journey	Source	Destination	Route	Dep_Time	Arrival_Time	Duration	Total_Stops	Additional_Info	Price
0	IndiGo	24/03/2019	Banglore	New Delhi	BLR → DEL	22:20	01:10 22 Mar	2h 50m	non-stop	No info	3897
1	Air India	1/05/2019	Kolkata	Banglore	CCU → IXR → BBI → BLR	05:50	13:15	7h 25m	2 stops	No info	7662

0s completed at 6:16 PM

Type here to search

ENG IN 6:17 PM 3/23/2023

Welcome To Colaboratory - Cola | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

0s

[2] 1s
```
data=pd.read_csv("/content/Data_Train_1.csv")
data.head()
```

	Airline	Date_of_Journey	Source	Destination	Route	Dep_Time	Arrival_Time	Duration	Total_Stops	Additional_Info	Price
0	IndiGo	24/03/2019	Banglore	New Delhi	BLR → DEL	22:20	01:10 22 Mar	2h 50m	non-stop	No info	3897
1	Air India	1/05/2019	Kolkata	Banglore	CCU → IXR → BBI → BLR	05:50	13:15	7h 25m	2 stops	No info	7662
2	Jet Airways	9/06/2019	Delhi	Cochin	DEL → LKO → BOM → COK	09:25	04:25 10 Jun	19h	2 stops	No info	13882
3	IndiGo	12/05/2019	Kolkata	Banglore	CCU → NAG → BLR	18:05	23:30	5h 25m	1 stop	No info	6218
4	IndiGo	01/03/2019	Banglore	New Delhi	BLR → NAG → DEL	16:50	21:35	4h 45m	1 stop	No info	13302

[3] 0s
```
data.Date_of_journey=data.Date_of_Journey.str.split('/')
data.Date_of_Journey
```

```
0        24/03/2019
1         1/05/2019
2         9/06/2019
3        12/05/2019
4        01/03/2019
            ...
10678     9/04/2019
10679    27/04/2019
10680    27/04/2019
```

0s completed at 6:16 PM

Type here to search

ENG IN 6:17 PM 3/23/2023

Welcome To Colaboratory - Cola | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

```
[3] data.Date_of_journey=data.Date_of_Journey.str.split('/')
    data.Date_of_Journey
```

```
0        24/03/2019
1         1/05/2019
2         9/06/2019
3        12/05/2019
4        01/03/2019
            ...
10678     9/04/2019
10679    27/04/2019
10680    27/04/2019
10681    01/03/2019
10682     9/05/2019
Name: Date_of_Journey, Length: 10683, dtype: object
```

```
[6]
    data['Date']=data.Date_of_Journey.str[0]
    data['Month']=data.Date_of_Journey.str[1]
    data['Year']=data.Date_of_Journey.str[2]
```

```
[7]
    data.Total_Stops.unique()
```

```
array(['non-stop', '2 stops', '1 stop', '3 stops', nan, '4 stops'],
      dtype=object)
```

0s completed at 6:16 PM

Type here to search

ENG IN 6:17 PM 3/23/2023

Welcome To Colaboratory - Colal | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

RAM Disk

+ Code + Text

[7]
```
data.Total_Stops.unique()
```

```
array(['non-stop', '2 stops', '1 stop', '3 stops', nan, '4 stops'],
      dtype=object)
```

[8]
```
data.Route=data.Route.str.split('')
data.Route
```

```
0                        [, B, L, R,  , →,  , D, E, L, ]
1        [, C, C, U,  , →,  , I, X, R,  , →,  , B, B, I...
2        [, D, E, L,  , →,  , L, K, O,  , →,  , B, O, M...
3        [, C, C, U,  , →,  , N, A, G,  , →,  , B, L, R, ]
4        [, B, L, R,  , →,  , N, A, G,  , →,  , D, E, L, ]
                                ...
10678                    [, C, C, U,  , →,  , B, L, R, ]
10679                    [, C, C, U,  , →,  , B, L, R, ]
10680                    [, B, L, R,  , →,  , D, E, L, ]
10681                    [, B, L, R,  , →,  , D, E, L, ]
10682    [, D, E, L,  , →,  , G, O, I,  , →,  , B, O, M...
Name: Route, Length: 10683, dtype: object
```

[ ]
```
data['City1']=data.Route.str[0]
data['City2']=data.Route.str[1]
data['City3']=data.Route.str[2]
data['City4']=data.Route.str[3]
```

0s completed at 6:16 PM

Type here to search

ENG IN 6:17 PM 3/23/2023

Welcome To Colaboratory - Cola | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

```
[9]
data.Dep_Time=data.Dep_Time.str.split(':')
```

```
[10] data['Dep_Time_Hour']=data.Dep_Time.str[0]
data['Dep_Time_Mins']=data.Dep_Time.str[1]
```

```
[11]
data.Arrival_Time=data.Arrival_Time.str.split('')
```

```
[12]
data['Arrival_date']=data.Arrival_Time.str[1]
data['Time_of_Arrival']=data.Arrival_Time.str[0]
```

```
[13] data['Time_of_Arrival']=data.Time_of_Arrival.str.split(':')
```

```
[14] data['Arrival_Time_Hour']=data.Time_of_Arrival.str[0]
data['Arrival_Time_Mins']=data.Time_of_Arrival.str[1]
```

```
[15]
data.Duration=data.Duration.str.split('')
```

```
[16] data['Travel_Hours']=data.Duration.str[0]
```

0s completed at 6:16 PM

Type here to search

ENG IN 6:17 PM 3/23/2023

Welcome To Colaboratory - Cola | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

```
[16] data.Travel_Mins=data.Travel_Mins.str.split('m')
     data.Travel_Mins=data.Travel_Mins.str[0]
```

```
[17] data.Total_Stops.replace('non_stop',0,inplace=True)
     data.Total_stops=data.Total_Stops.str.split('')
     data.Total_Stops=data.Total_Stops.str[0]
```

```
[18] data.Additional_Info.unique()
```

```
array(['No info', 'In-flight meal not included',
       'No check-in baggage included', '1 Short layover', 'No Info',
       '1 Long layover', 'Change airports', 'Business class',
       'Red-eye flight', '2 Long layover'], dtype=object)
```

```
[19] data.Additional_Info.replace('No Info','No info',inplace=True)
```

```
[20] data.isnull().sum()
```

```
Airline            0
Date_of_Journey    0
Source             0
Destination        0
Route              1
```

0s completed at 6:16 PM

Type here to search

ENG IN 6:18 PM 3/23/2023

Welcome To Colaboratory - Cola | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

[20] data.isnull().sum()

0s

```
Airline                0
Date_of_Journey        0
Source                 0
Destination            0
Route                  1
Dep_Time               0
Arrival_Time           0
Duration               0
Total_Stops            1
Additional_Info        0
Price                  0
Date                   0
Month                  0
Year                   0
Dep_Time_Hour          0
Dep_Time_Mins          0
Arrival_date           0
Time_of_Arrival        0
Arrival_Time_Hour      0
Arrival_Time_Mins    10683
Travel_Hours           0
Travel_Mins            0
dtype: int64
```

[21] data.drop(['City4','City5','City6'],axis=1,inplace=True)

0s completed at 6:16 PM

Type here to search

ENG IN 6:18 PM 3/23/2023

Welcome To Colaboratory - Cola | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

[21]

[22] 0s
```
data.drop(['Date_of_Journey','Route','Dep_Time','Arrival_Time','Duration'],axis=1,inplace=True)
data.drop(['Time_of_Arrival'],axis=1,inplace=True)
```

[23] 0s
```
data.isnull().sum()
```

```
Airline              0
Source               0
Destination          0
Total_Stops          1
Additional_Info      0
Price                0
Date                 0
Month                0
Year                 0
Dep_Time_Hour        0
Dep_Time_Mins        0
Arrival_date         0
Arrival_Time_Hour    0
Arrival_Time_Mins    10683
Travel_Hours         0
Travel_Mins          0
dtype: int64
```

0s completed at 6:16 PM

Type here to search

ENG IN 6:18 PM 3/23/2023

Welcome To Colaboratory - Cola | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

```
[25] data['Arrival_date'].fillna(data['Date'],inplace=True)
```

```
[26] data['Travel_Mins'].fillna(0,inplace=True)
```

```
[27] data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10683 entries, 0 to 10682
Data columns (total 16 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Airline           10683 non-null  object
 1   Source            10683 non-null  object
 2   Destination       10683 non-null  object
 3   Total_Stops       10682 non-null  object
 4   Additional_Info   10683 non-null  object
 5   Price             10683 non-null  int64
 6   Date              10683 non-null  object
 7   Month             10683 non-null  object
 8   Year              10683 non-null  object
 9   Dep_Time_Hour     10683 non-null  object
 10  Dep_Time_Mins     10683 non-null  object
 11  Arrival_date      10683 non-null  object
 12  Arrival_Time_Hour 10683 non-null  object
 13  Arrival_Time_Mins 0 non-null      float64
 14  Travel_Hours      10683 non-null  object
```

0s completed at 6:16 PM

Type here to search

ENG IN 6:18 PM 3/23/2023

Welcome To Colaboratory - Colal | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

```
[28] data.Date=data.Date.astype('int64')
     data.Month=data.Month.astype('int64')
     data.Year=data.Year.astype('astype')
     data.Dep_Time_Hour=data.Dep_Time_Hour.astype('astype')
     data.Dep_Time_Hour=data.Dep_Time_Hour.astype('int64')
     data.Dep_Time_Mins=data.Dep_Time_Mins.astype('int64')
     data.Arrival_date=data.Arrival_date.astype("int64")
     data.Arrival_Time_Hour=data.Arrival_Time_Hour.astype('int64')
     data.Arrival_Time_Mins=data.Arrival_Time_Mins.astype('int64')
     data.Travel_Mins=data.Travel_Mins.astype('int64')
```

```
[29] data[data['Travel_Hours']=='5m']
```

0s

	Airline	Source	Destination	Total_Stops	Additional_Info	Price	Date	Month	Year	Dep_Time_Hour	Dep_Time_Mins	Arrival_date	Arrival_Time_Hour	Arrival_Time

```
[30] data.drop(index=6474,inplace=True,axis=0)
```

0s

```
[31] data.Travel_Hours=data.Travel_Hours.astype('int64')
```

0s completed at 6:16 PM

Type here to search

ENG IN 6:18 PM 3/23/2023

Welcome To Colaboratory - Cola | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

```
[30] data.drop(index=6474,inplace=True,axis=0)
```

```
[31] data.Travel_Hours=data.Travel_Hours.astype('int64')
```

```
[32] categorical=['Airline','Source','Destination','Additional_Info','City1']
     numerical=['Total_Stops','Date','Month','Year','Dep_Time_Hour','Dep_Time_Mins','Arrival_Time_Hour','Arrival_Time_Mins','Travel_Hours','Travel_Mins']
```

```
[33] from sklearn.preprocessing import LabelEncoder
     le=LabelEncoder()
```

```
[34] data.Airline=le.fit_transform(data.Airline)
     data.Source=le.fit_transform(data.Source)
     data.Destination=le.fit_transform(data.Destination)
     data.Total_Stops=le.fit_transform(data.Total_Stops)
     data.City1    =le.fit_transform(data.City1    )
     data.City2=le.fit_transform(data.City2)
     data.City3=le.fit_transform(data.City3)
     data.Additional_Info=le.fit_transform(data.Additional_Info)
     data.head()
```

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

0s completed at 6:16 PM

Type here to search

ENG IN 6:18 PM 3/23/2023

Welcome To Colaboratory - Colal | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

[35] data.head()

	Airline	Source	Destination	Total_Stops	Additional_Info	Price	Date	Month	Year	Dep_Time_Hour	Dep_Time_Mins	Arrival_date	Arrival_Time_Hour	Arrival_Tim
0	3	0	5	4	No info	3897	2	4	/	22	20	0		
1	1	3	0	1	No info	7662	1	/	0	05	50	1		
2	4	2	1	1	No info	13882	9	/	0	09	25	0		
3	3	3	0	0	No info	6218	1	2	/	18	05	2		
4	3	0	5	0	No info	13302	0	1	/	16	50	2		

```
data=data[['Airline','Source','Destination','Date','Month','Year','Dep_Time_Mins','Arrival_date','Arrival_Time']]
```

[37] data.head()

	Airline	Source	Destination	Total_Stops	Additional_Info	Price	Date	Month	Year	Dep_Time_Hour	Dep_Time_Mins	Arrival_date	Arrival_Time_Hour	Arrival_Tim
0	3	0	5	4	No info	3897	2	4	/	22	20	0		

0s completed at 6:16 PM

Type here to search

ENG IN 6:18 PM 3/23/2023

Welcome To Colaboratory - Colal | Untitled5.ipynb - Colaboratory

colab.research.google.com/drive/14NwujhZWMvbEi7NAFDFhDRZVsr8YIjm1#scrollTo=7v7MN680MAYv

Untitled5.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk

[35]

3	3	3	0	0	No info	6218	1	2	/	18	05	2
4	3	0	5	0	No info	13302	0	1	/	16	50	2

```
data=data[['Airline','Source','Destination','Date','Month','Year','Dep_Time_Mins','Arrival_date','Arrival_Time']]
```

[37]

```
data.head()
```

	Airline	Source	Destination	Total_Stops	Additional_Info	Price	Date	Month	Year	Dep_Time_Hour	Dep_Time_Mins	Arrival_date	Arrival_Time_Hour	Arrival_Tim
0	3	0	5	4	No info	3897	2	4	/	22	20	0		
1	1	3	0	1	No info	7662	1	/	0	05	50	1		
2	4	2	1	1	No info	13882	9	/	0	09	25	0		
3	3	3	0	0	No info	6218	1	2	/	18	05	2		
4	3	0	5	0	No info	13302	0	1	/	16	50	2		

0s completed at 6:16 PM

Type here to search

ENG IN 6:19 PM 3/23/2023