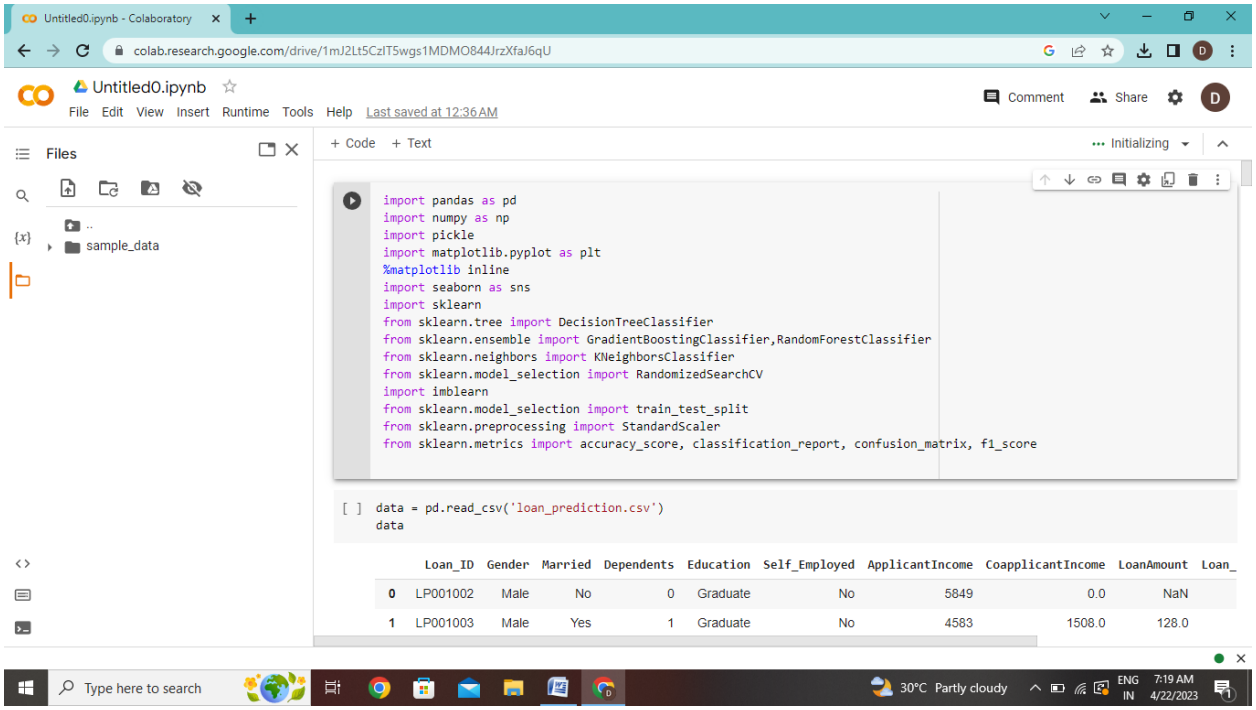
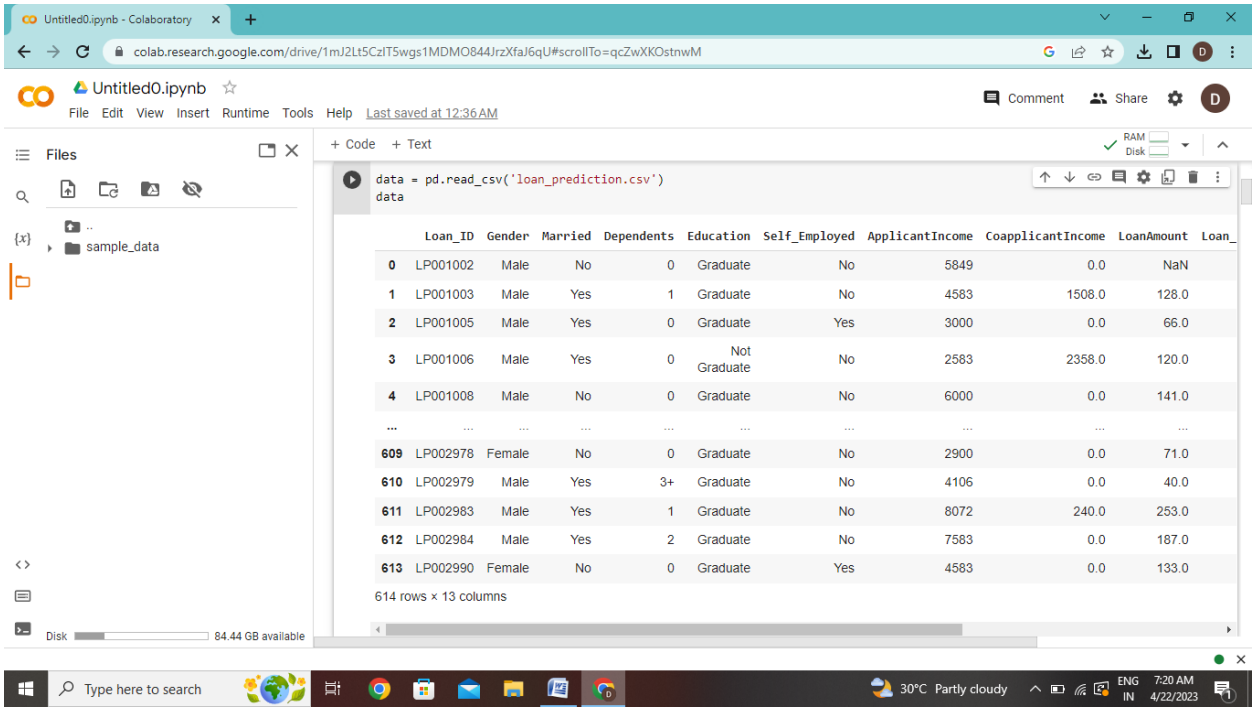
OUTPUT


Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
+ Code + Text
Initializing
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.tree import DecisionTreeClassifier
from sklearn.ensemble import GradientBoostingClassifier,RandomForestClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import RandomizedSearchCV
import imblearn
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
[ ] data = pd.read_csv('loan_prediction.csv')
data
Loan_ID Gender Married Dependents Education Self_Employed ApplicantIncome CoapplicantIncome LoanAmount Loan_
0 LP001002 Male No 0 Graduate No 5849 0.0 NaN
1 LP001003 Male Yes 1 Graduate No 4583 1508.0 128.0
Type here to search
30°C Partly cloudy
ENG IN
7:19 AM 4/22/2023



Untitled0.ipynb - Colaboratory

colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=qcZwXKOstnwM

Untitled0.ipynb

File Edit View Insert Runtime Tools Help Last saved at 12:36 AM

Comment Share

Files

sample_data

Disk 84.44 GB available

+ Code + Text

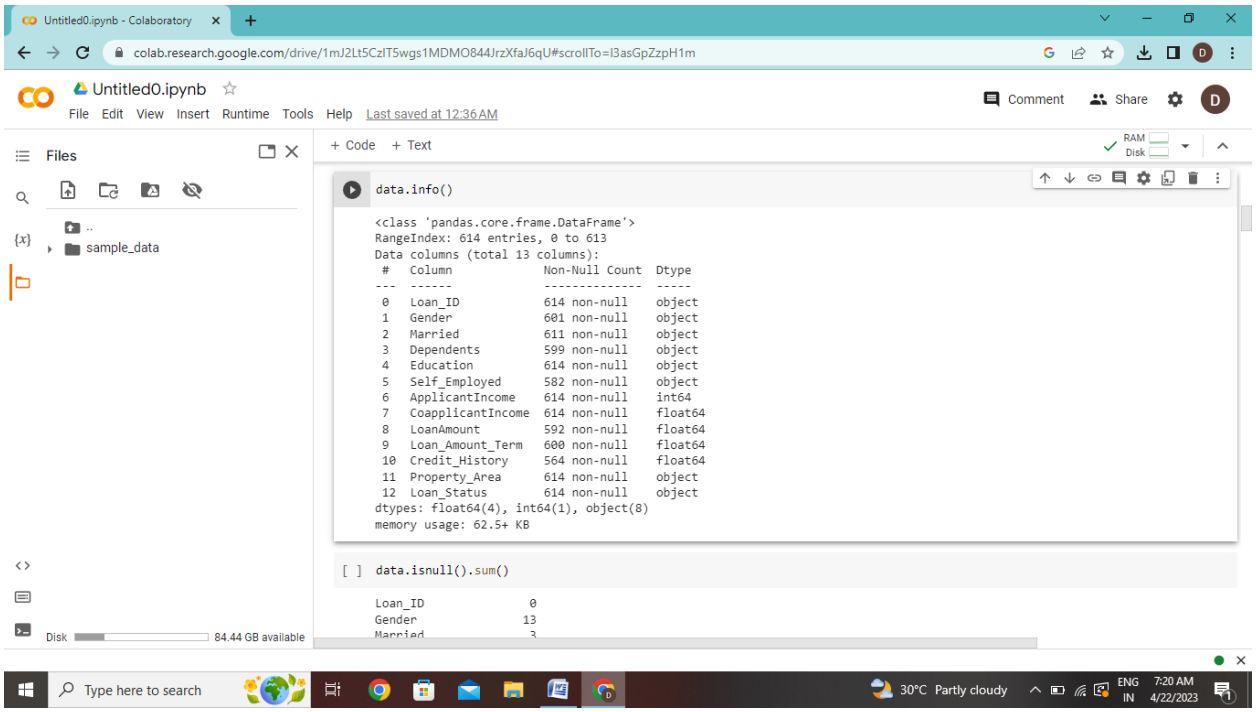
RAM Disk

```
data = pd.read_csv('loan_prediction.csv')
data
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_
0	LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120.0	
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	
...	...	...	...	...	...	...	...	...	...	
609	LP002978	Female	No	0	Graduate	No	2900	0.0	71.0	
610	LP002979	Male	Yes	3+	Graduate	No	4106	0.0	40.0	
611	LP002983	Male	Yes	1	Graduate	No	8072	240.0	253.0	
612	LP002984	Male	Yes	2	Graduate	No	7583	0.0	187.0	
613	LP002990	Female	No	0	Graduate	Yes	4583	0.0	133.0	

614 rows × 13 columns

Type here to search 30°C Partly cloudy ENG IN 7:20 AM 4/22/2023



Untitled0.ipynb - Colaboratory

colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=l3asGpZzpH1m

Untitled0.ipynb

File Edit View Insert Runtime Tools Help Last saved at 12:36 AM

Comment Share

Files

sample_data

Disk 84.44 GB available

+ Code + Text

RAM Disk

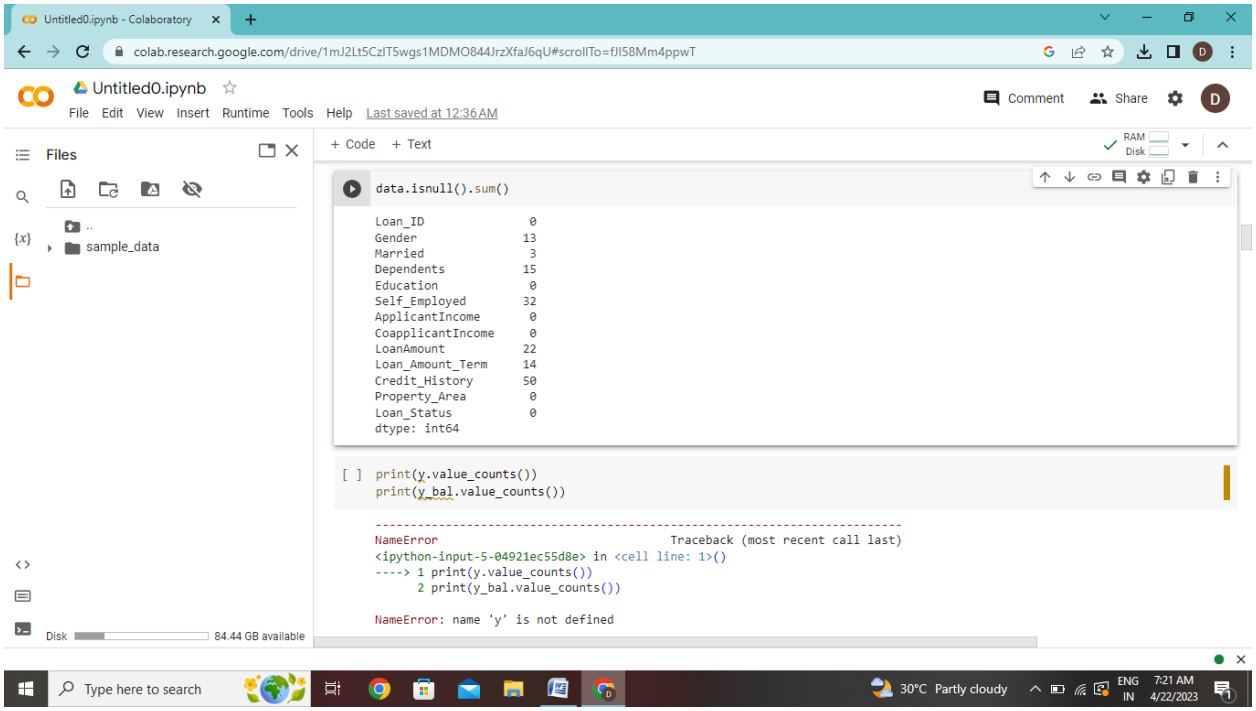
```
data.info()
```

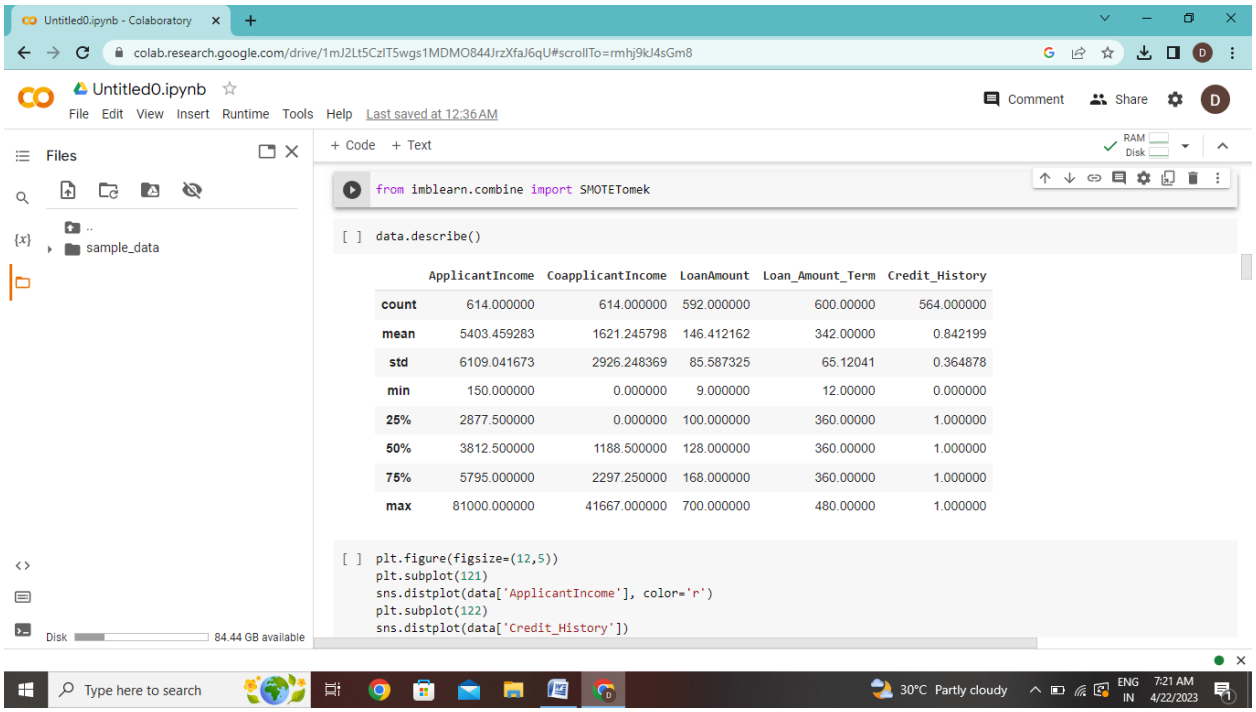
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 614 entries, 0 to 613
Data columns (total 13 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   Loan_ID            614 non-null    object
 1   Gender             601 non-null    object
 2   Married            611 non-null    object
 3   Dependents         599 non-null    object
 4   Education          614 non-null    object
 5   Self_Employed      582 non-null    object
 6   ApplicantIncome    614 non-null    int64
 7   CoapplicantIncome  614 non-null    float64
 8   LoanAmount         592 non-null    float64
 9   Loan_Amount_Term   600 non-null    float64
 10  Credit_History     564 non-null    float64
 11  Property_Area      614 non-null    object
 12  Loan_Status        614 non-null    object
dtypes: float64(4), int64(1), object(8)
memory usage: 62.5+ KB
```

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

```
Loan_ID             0
Gender             13
Married             3
```

Type here to search 30°C Partly cloudy ENG IN 7:20 AM 4/22/2023


Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=fJI58Mm4ppwT
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
Disk 84.44 GB available
+ Code + Text
RAM
Disk
data.isnull().sum()
Loan_ID 0
Gender 13
Married 3
Dependents 15
Education 0
Self_Employed 32
ApplicantIncome 0
CoapplicantIncome 0
LoanAmount 22
Loan_Amount_Term 14
Credit_History 50
Property_Area 0
Loan_Status 0
dtype: int64
[ ] print(y.value_counts())
print(y_bal.value_counts())
NameError Traceback (most recent call last)
<ipython-input-5-04921ec55d8e> in <cell line: 1>()
----> 1 print(y.value_counts())
2 print(y_bal.value_counts())
NameError: name 'y' is not defined
Type here to search
30°C Partly cloudy
ENG IN
7:21 AM
4/22/2023


Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=rmhj9kJ4sGm8
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
Disk 84.44 GB available
+ Code + Text
RAM
Disk
from imblearn.combine import SMOTETomek
[ ] data.describe()
ApplicantIncome CoapplicantIncome LoanAmount Loan_Amount_Term Credit_History
count 614.000000 614.000000 592.000000 600.00000 564.000000
mean 5403.459283 1621.245798 146.412162 342.00000 0.842199
std 6109.041673 2926.248369 85.587325 65.12041 0.364878
min 150.000000 0.000000 9.000000 12.00000 0.000000
25% 2877.500000 0.000000 100.000000 360.00000 1.000000
50% 3812.500000 1188.500000 128.000000 360.00000 1.000000
75% 5795.000000 2297.250000 168.000000 360.00000 1.000000
max 81000.000000 41667.000000 700.000000 480.00000 1.000000
[ ] plt.figure(figsize=(12,5))
plt.subplot(121)
sns.distplot(data['ApplicantIncome'], color='r')
plt.subplot(122)
sns.distplot(data['Credit_History'])
Type here to search
30°C Partly cloudy
ENG IN
7:21 AM
4/22/2023

Untitled0.ipynb - Colaboratory

colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=9-knz--htL_M

Untitled0.ipynb

File Edit View Insert Runtime Tools Help Last saved at 12:36 AM

Comment Share

Files

sample_data

+ Code + Text

RAM Disk

```
plt.figure(figsize=(12,5))
plt.subplot(121)
sns.distplot(data['ApplicantIncome'], color='r')
plt.subplot(122)
sns.distplot(data['Credit_History'])
plt.show()
```

```
<ipython-input-13-4b78f43a4171>:3: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(data['ApplicantIncome'], color='r')
<ipython-input-13-4b78f43a4171>:5: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(data['Credit_History'])
```

Disk 84.44 GB available

Type here to search

30°C Partly cloudy ENG IN 7:22 AM 4/22/2023

Untitled0.ipynb - Colaboratory

colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=9-knz--htL_M

Untitled0.ipynb

File Edit View Insert Runtime Tools Help Last saved at 12:36 AM

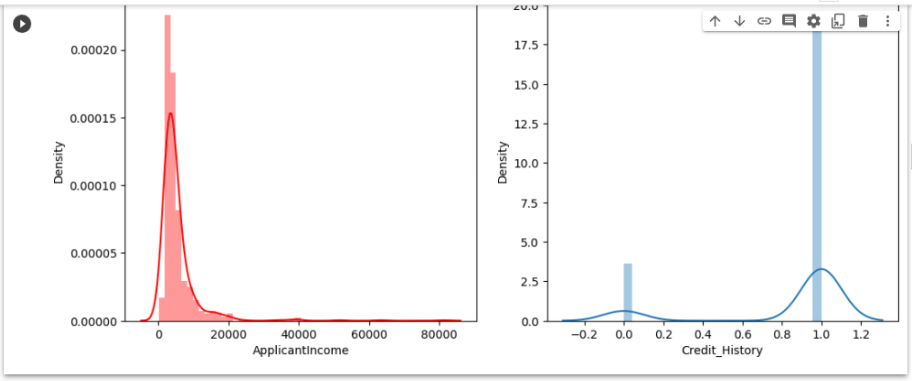
Comment Share

Files

sample_data

+ Code + Text

RAM Disk

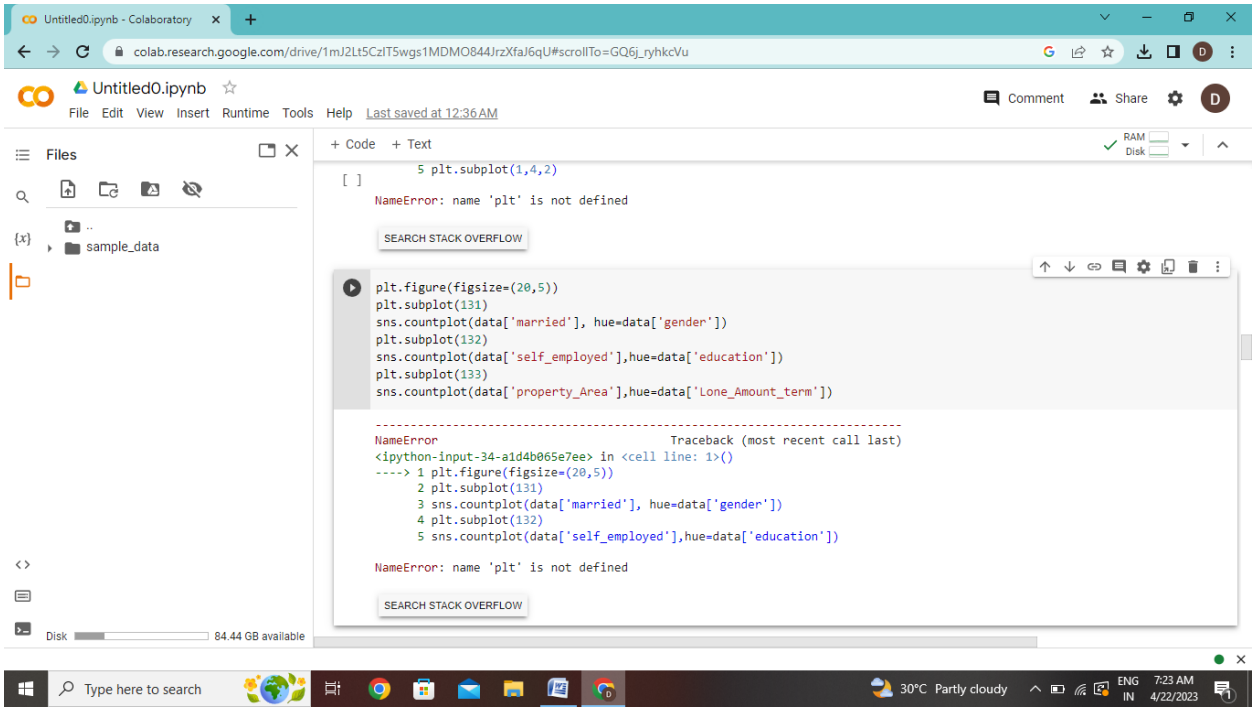


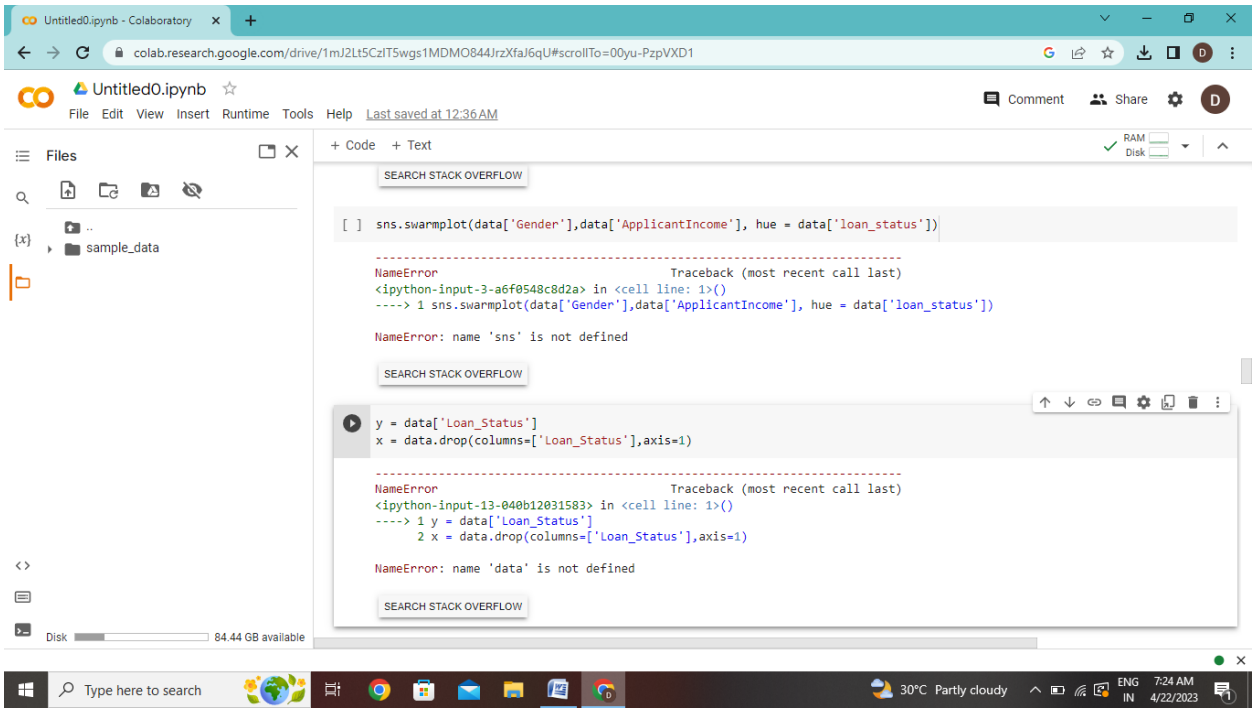

```
[ ] #plotting the count plot
    plt.figure(figsize=(18,4))
    plt.subplot(1,4,1)
    sns.countplot(data['Gender'])
    plt.subplot(1,4,2)
```

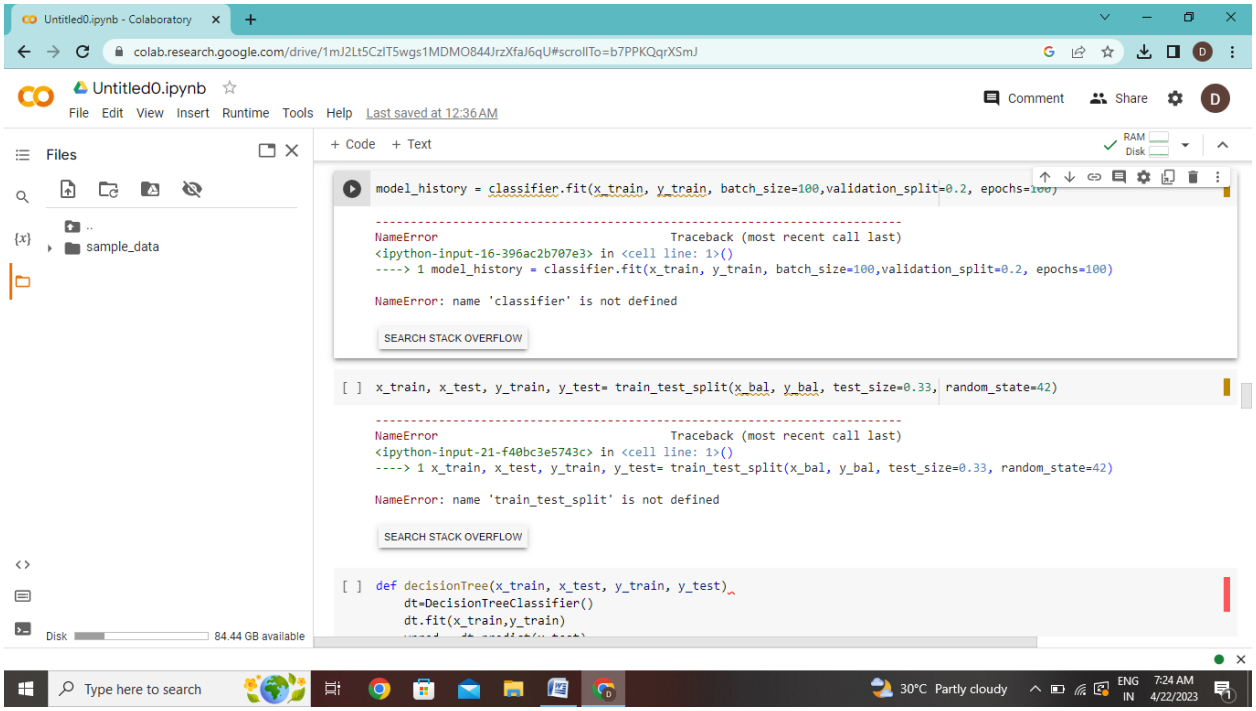
Disk 84.44 GB available

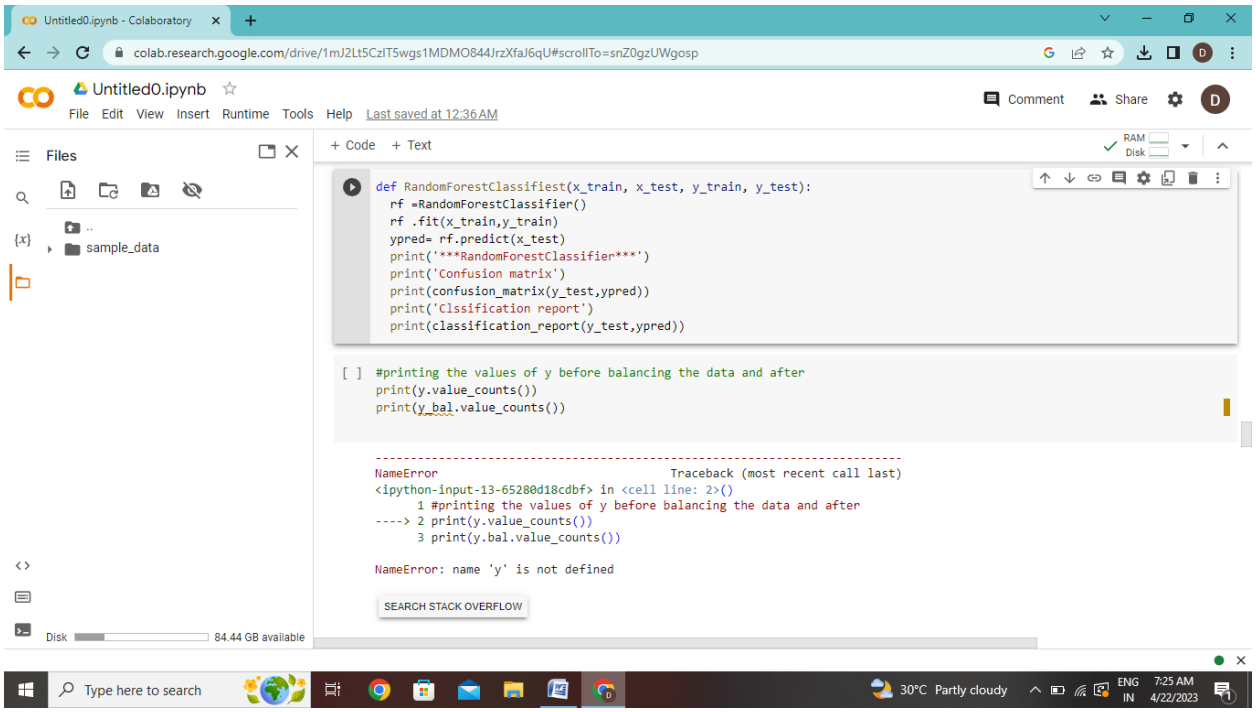
Type here to search

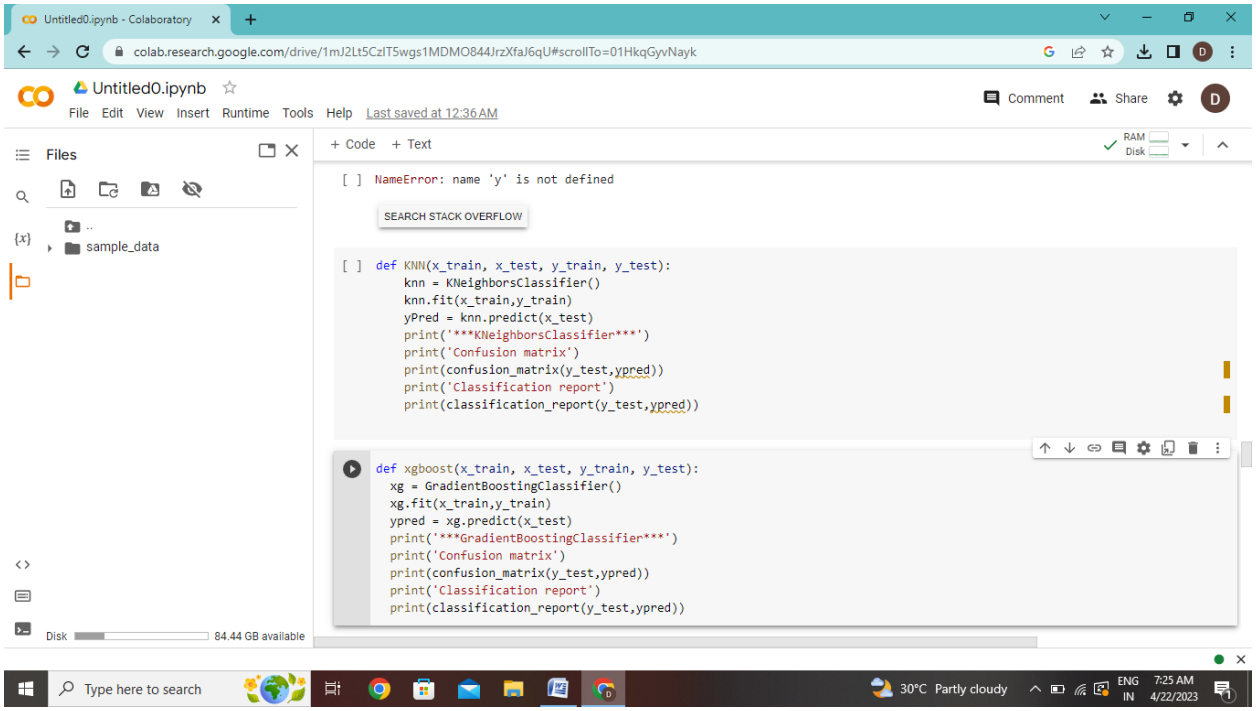
30°C Partly cloudy ENG IN 7:23 AM 4/22/2023

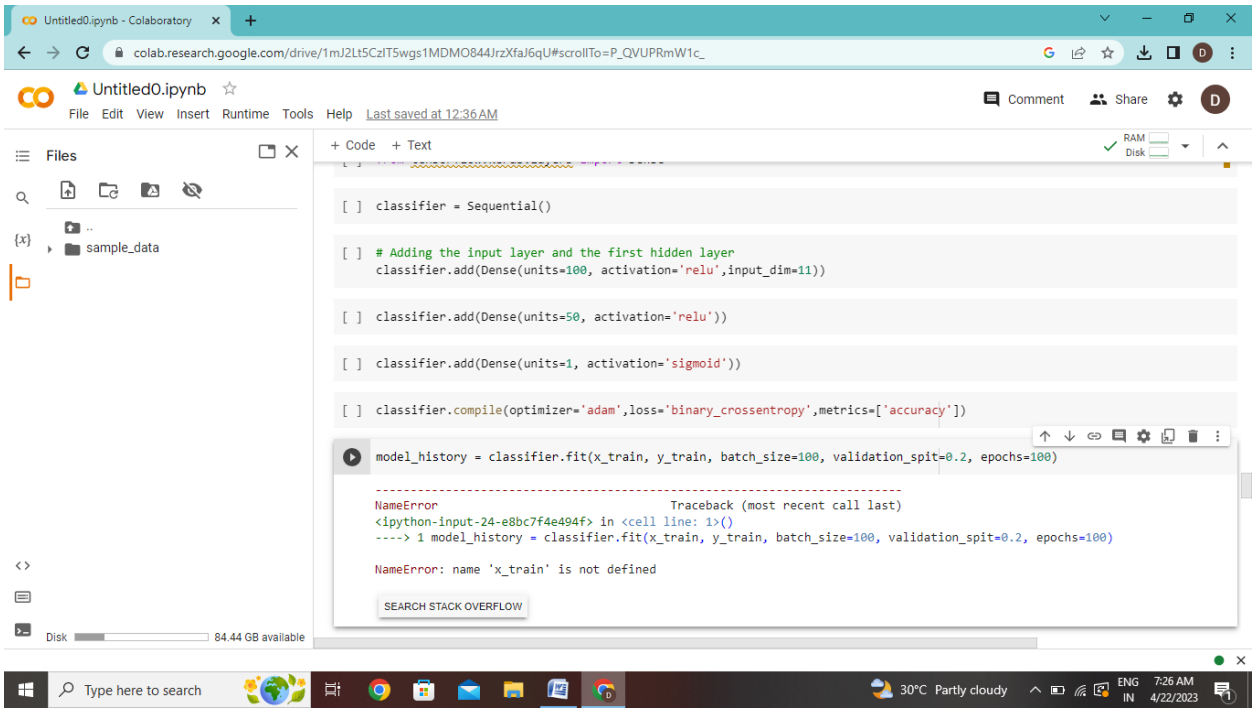

Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=GQ6j_ryhkcVu
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
+ Code + Text
RAM
Disk
5 plt.subplot(1,4,2)
NameError: name 'plt' is not defined
SEARCH STACK OVERFLOW
plt.figure(figsize=(20,5))
plt.subplot(131)
sns.countplot(data['married'], hue=data['gender'])
plt.subplot(132)
sns.countplot(data['self_employed'],hue=data['education'])
plt.subplot(133)
sns.countplot(data['property_Area'],hue=data['Lone_Amount_term'])
NameError Traceback (most recent call last)
<ipython-input-34-a1d4b065e7ee> in <cell line: 1>()
----> 1 plt.figure(figsize=(20,5))
2 plt.subplot(131)
3 sns.countplot(data['married'], hue=data['gender'])
4 plt.subplot(132)
5 sns.countplot(data['self_employed'],hue=data['education'])
NameError: name 'plt' is not defined
SEARCH STACK OVERFLOW
Disk 84.44 GB available
Type here to search
30°C Partly cloudy
ENG IN 7:23 AM 4/22/2023

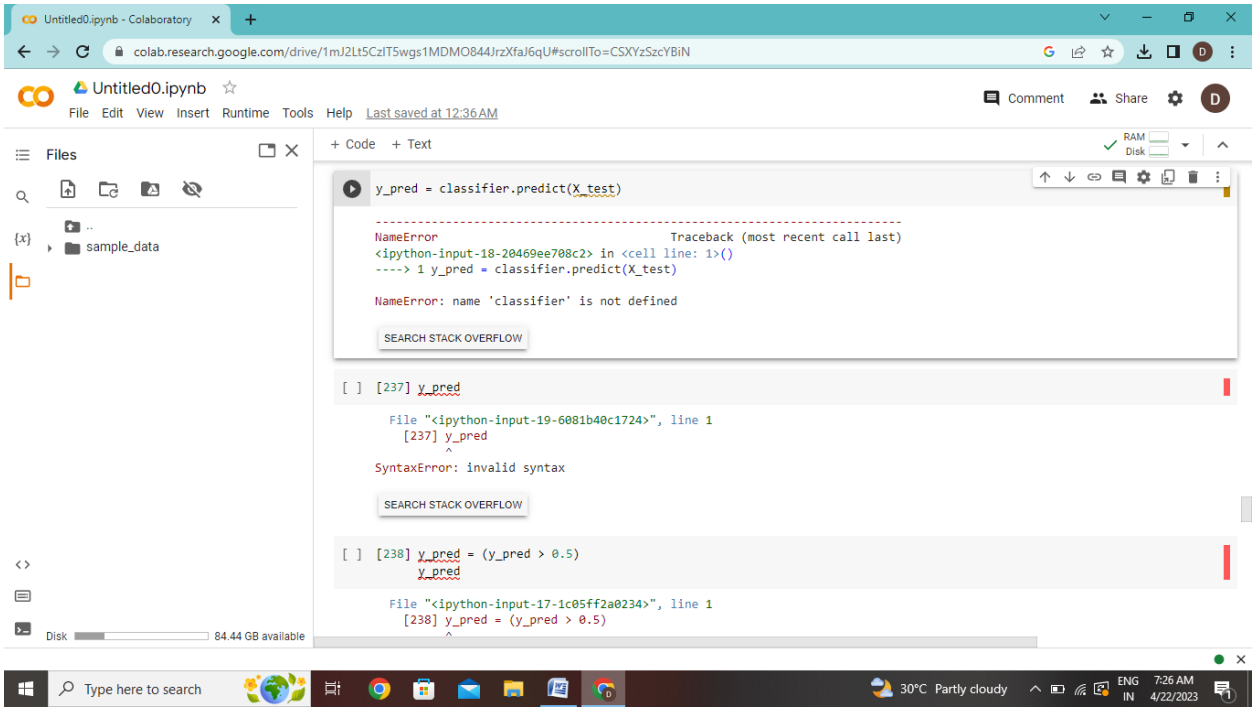

Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=00yu-PzpVXD1
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
+ Code + Text
RAM
Disk
SEARCH STACK OVERFLOW
sns.swarmplot(data['Gender'],data['ApplicantIncome'], hue = data['loan_status'])
NameError Traceback (most recent call last)
<ipython-input-3-a6f0548c8d2a> in <cell line: 1>()
----> 1 sns.swarmplot(data['Gender'],data['ApplicantIncome'], hue = data['loan_status'])
NameError: name 'sns' is not defined
SEARCH STACK OVERFLOW
y = data['Loan_Status']
x = data.drop(columns=['Loan_Status'],axis=1)
NameError Traceback (most recent call last)
<ipython-input-13-040b12031583> in <cell line: 1>()
----> 1 y = data['Loan_Status']
2 x = data.drop(columns=['Loan_Status'],axis=1)
NameError: name 'data' is not defined
SEARCH STACK OVERFLOW
Disk 84.44 GB available
Type here to search
30°C Partly cloudy
ENG IN 7:24 AM 4/22/2023

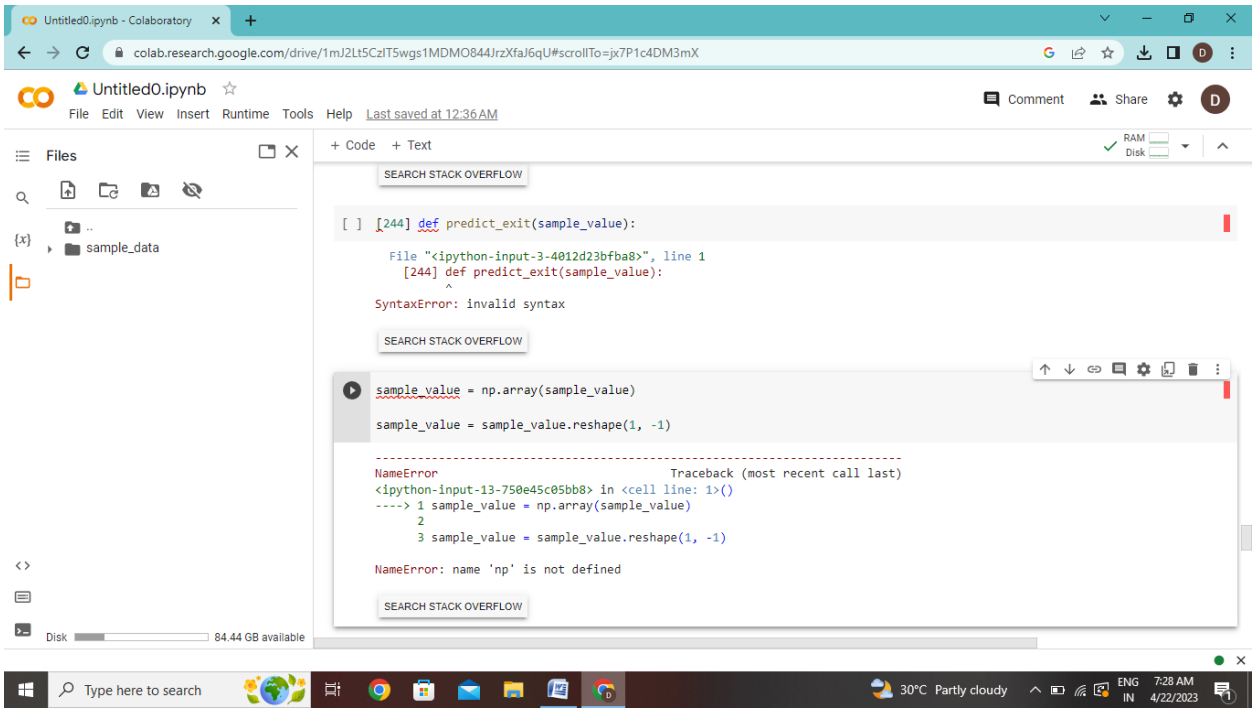

Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=b7PPKQqrXSmJ
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
+ Code + Text
RAM
Disk
model_history = classifier.fit(x_train, y_train, batch_size=100,validation_split=0.2, epochs=100)
NameError Traceback (most recent call last)
<ipython-input-16-396ac2b707e3> in <cell line: 1>()
----> 1 model_history = classifier.fit(x_train, y_train, batch_size=100,validation_split=0.2, epochs=100)
NameError: name 'classifier' is not defined
SEARCH STACK OVERFLOW
[ ] x_train, x_test, y_train, y_test= train_test_split(x_bal, y_bal, test_size=0.33, random_state=42)
NameError Traceback (most recent call last)
<ipython-input-21-f40bc3e5743c> in <cell line: 1>()
----> 1 x_train, x_test, y_train, y_test= train_test_split(x_bal, y_bal, test_size=0.33, random_state=42)
NameError: name 'train_test_split' is not defined
SEARCH STACK OVERFLOW
[ ] def decisionTree(x_train, x_test, y_train, y_test)
dt=DecisionTreeClassifier()
dt.fit(x_train,y_train)
Disk 84.44 GB available
Type here to search
30°C Partly cloudy
ENG IN 7:24 AM 4/22/2023

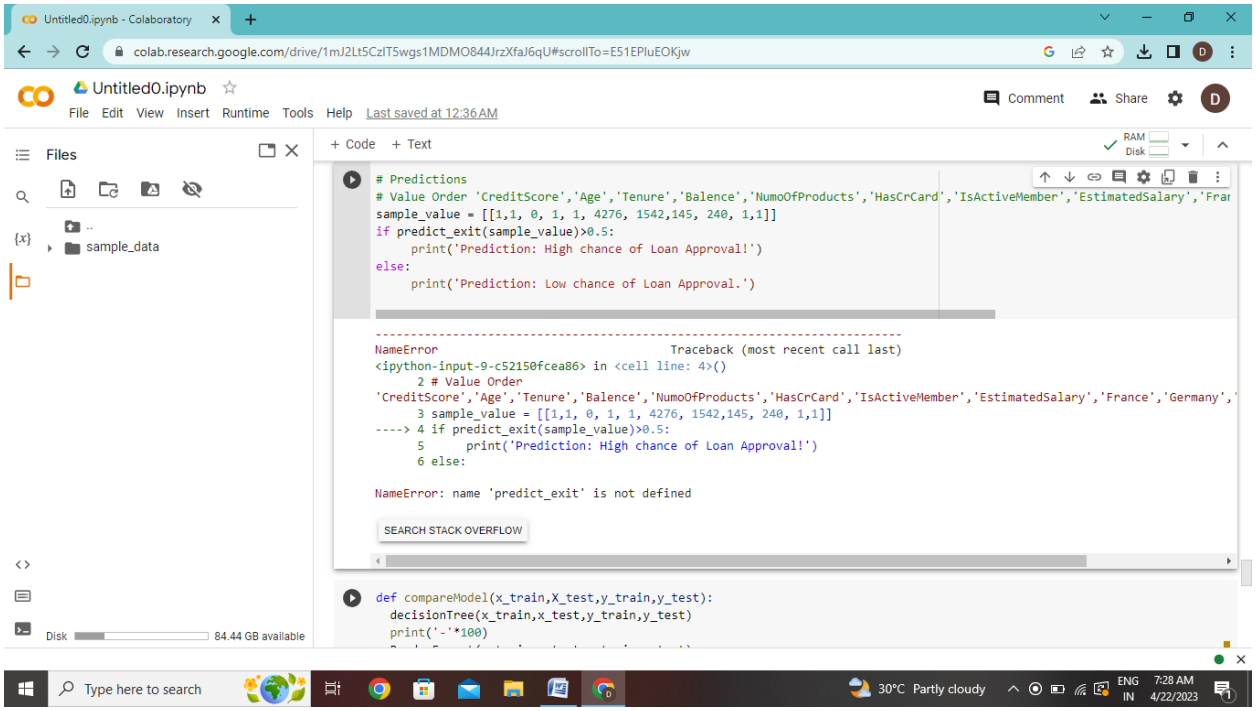

Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=snZ0gzUWgosp
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
+ Code + Text
RAM
Disk
def RandomForestClassifiest(x_train, x_test, y_train, y_test):
rf =RandomForestClassifier()
rf .fit(x_train,y_train)
ypred= rf.predict(x_test)
print('***RandomForestClassifier***')
print('Confusion matrix')
print(confusion_matrix(y_test,ypred))
print('Clssification report')
print(classification_report(y_test,ypred))
[ ] #printing the values of y before balancing the data and after
print(y.value_counts())
print(y_bal.value_counts())
NameError Traceback (most recent call last)
<ipython-input-13-65280d18cdbf> in <cell line: 2>()
1 #printing the values of y before balancing the data and after
----> 2 print(y.value_counts())
3 print(y.bal.value_counts())
NameError: name 'y' is not defined
SEARCH STACK OVERFLOW
Disk 84.44 GB available
Type here to search
30°C Partly cloudy
ENG IN 7:25 AM 4/22/2023

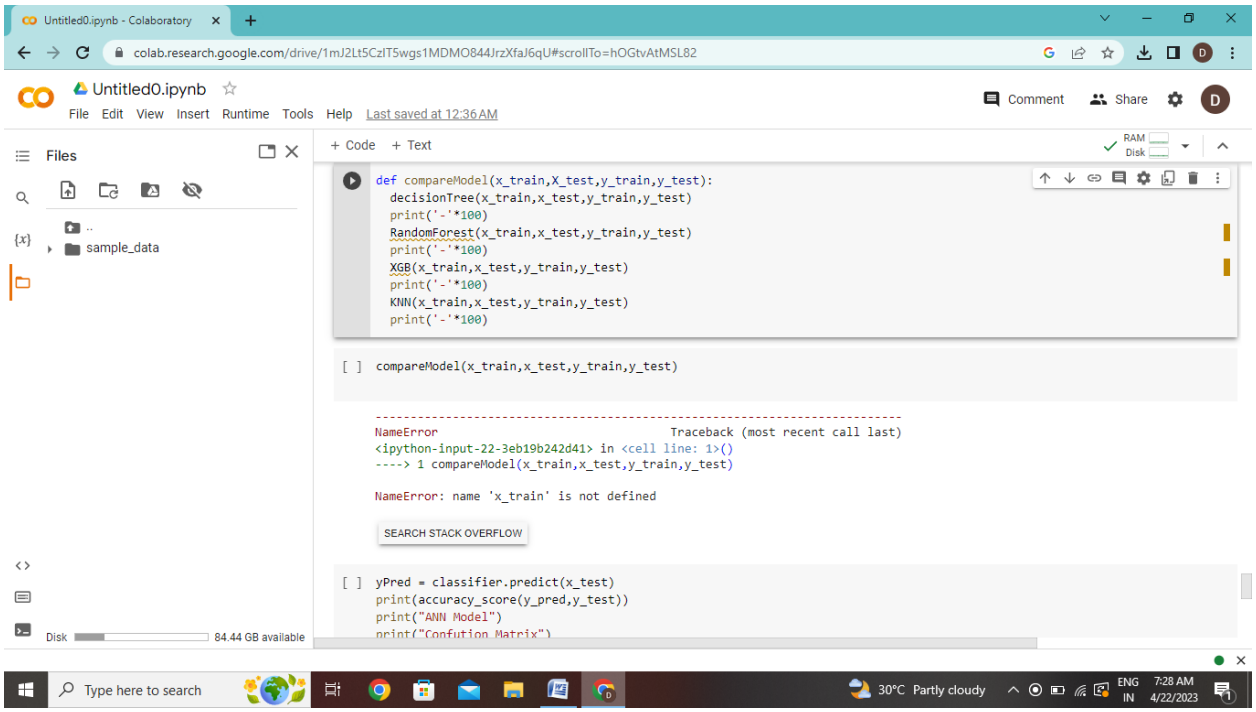

Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=01HkqGyvNayk
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
Disk 84.44 GB available
+ Code + Text
RAM
Disk
[ ] NameError: name 'y' is not defined
SEARCH STACK OVERFLOW
[ ] def KNN(x_train, x_test, y_train, y_test):
knn = KNeighborsClassifier()
knn.fit(x_train,y_train)
yPred = knn.predict(x_test)
print('***KNeighborsClassifier***')
print('Confusion matrix')
print(confusion_matrix(y_test,ypred))
print('Classification report')
print(classification_report(y_test,ypred))
def xgboost(x_train, x_test, y_train, y_test):
xg = GradientBoostingClassifier()
xg.fit(x_train,y_train)
ypred = xg.predict(x_test)
print('***GradientBoostingClassifier***')
print('Confusion matrix')
print(confusion_matrix(y_test,ypred))
print('Classification report')
print(classification_report(y_test,ypred))
Type here to search
30°C Partly cloudy
ENG IN
7:25 AM
4/22/2023

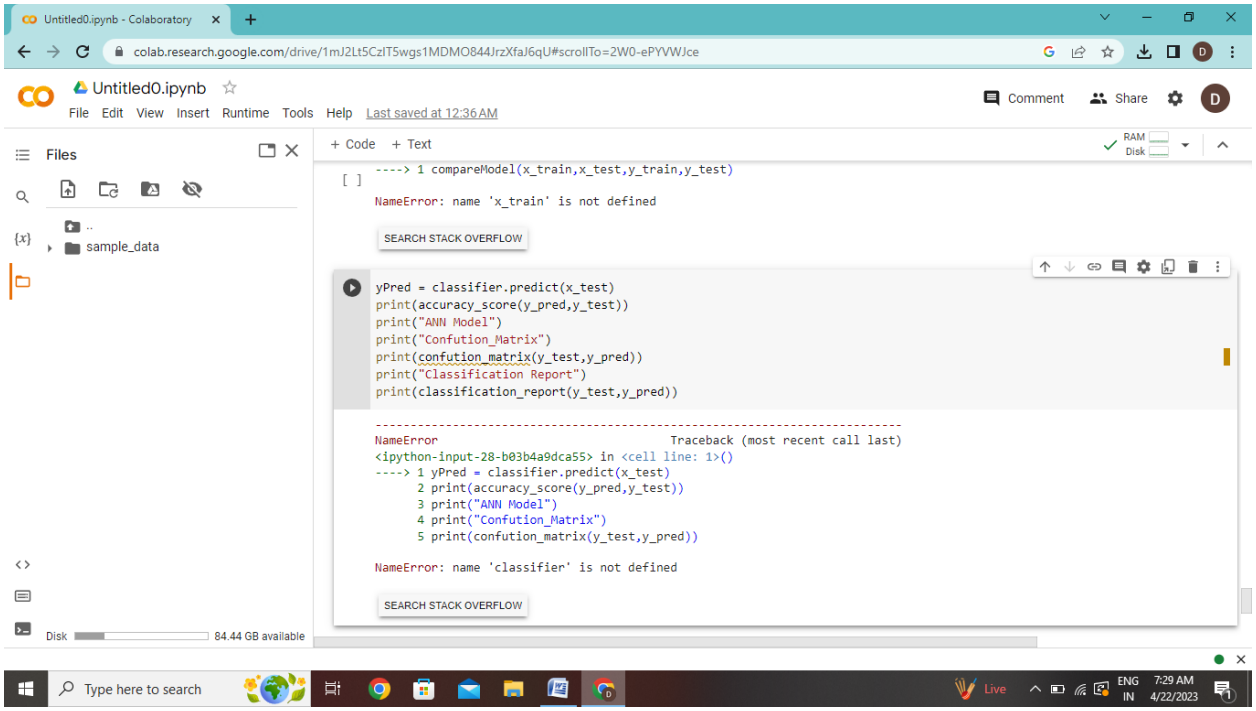

Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=P_QVUPRmW1c_
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
Disk 84.44 GB available
+ Code + Text
RAM
Disk
[ ] classifier = Sequential()
[ ] # Adding the input layer and the first hidden layer
classifier.add(Dense(units=100, activation='relu',input_dim=11))
[ ] classifier.add(Dense(units=50, activation='relu'))
[ ] classifier.add(Dense(units=1, activation='sigmoid'))
[ ] classifier.compile(optimizer='adam',loss='binary_crossentropy',metrics=['accuracy'])
model_history = classifier.fit(x_train, y_train, batch_size=100, validation_spit=0.2, epochs=100)
NameError Traceback (most recent call last)
<ipython-input-24-e8bc7f4e494f> in <cell line: 1>()
----> 1 model_history = classifier.fit(x_train, y_train, batch_size=100, validation_spit=0.2, epochs=100)
NameError: name 'x_train' is not defined
SEARCH STACK OVERFLOW
Type here to search
30°C Partly cloudy
ENG IN
7:26 AM
4/22/2023


Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=CSXYzSzcYBiN
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
y_pred = classifier.predict(X_test)
NameError Traceback (most recent call last)
<ipython-input-18-20469ee708c2> in <cell line: 1>()
----> 1 y_pred = classifier.predict(X_test)
NameError: name 'classifier' is not defined
SEARCH STACK OVERFLOW
[237] y_pred
File "<ipython-input-19-6081b40c1724>", line 1
[237] y_pred
^
SyntaxError: invalid syntax
SEARCH STACK OVERFLOW
[238] y_pred = (y_pred > 0.5)
y_pred
File "<ipython-input-17-1c05ff2a0234>", line 1
[238] y_pred = (y_pred > 0.5)
^
Disk
84.44 GB available
Type here to search
30°C Partly cloudy
ENG
IN
7:26 AM
4/22/2023


Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=jx7P1c4DM3mX
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
SEARCH STACK OVERFLOW
sample_data
[244] def predict_exit(sample_value):
File "<ipython-input-3-4012d23bfba8>", line 1
[244] def predict_exit(sample_value):
^
SyntaxError: invalid syntax
SEARCH STACK OVERFLOW
sample_value = np.array(sample_value)
sample_value = sample_value.reshape(1, -1)
NameError Traceback (most recent call last)
<ipython-input-13-750e45c05bb8> in <cell line: 1>()
----> 1 sample_value = np.array(sample_value)
2
3 sample_value = sample_value.reshape(1, -1)
NameError: name 'np' is not defined
SEARCH STACK OVERFLOW
Disk
84.44 GB available
Type here to search
30°C Partly cloudy
ENG
IN
7:28 AM
4/22/2023


Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=E51EPluEOKjw
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
Disk 84.44 GB available
+ Code + Text
RAM
Disk
# Predictions
# Value Order 'CreditScore','Age','Tenure','Balence','NumoOfProducts','HasCrCard','IsActiveMember','EstimatedSalary','Frar
sample_value = [[1,1, 0, 1, 1, 4276, 1542,145, 240, 1,1]]
if predict_exit(sample_value)>0.5:
print('Prediction: High chance of Loan Approval!')
else:
print('Prediction: Low chance of Loan Approval.')
NameError Traceback (most recent call last)
<ipython-input-9-c52150fcea86> in <cell line: 4>()
2 # Value Order
'CreditScore','Age','Tenure','Balence','NumoOfProducts','HasCrCard','IsActiveMember','EstimatedSalary','France','Germany','
3 sample_value = [[1,1, 0, 1, 1, 4276, 1542,145, 240, 1,1]]
----> 4 if predict_exit(sample_value)>0.5:
5 print('Prediction: High chance of Loan Approval!')
6 else:
NameError: name 'predict_exit' is not defined
SEARCH STACK OVERFLOW
def compareModel(x_train,X_test,y_train,y_test):
decisionTree(x_train,x_test,y_train,y_test)
print('-'*100)
Type here to search
30°C Partly cloudy
ENG IN 7:28 AM 4/22/2023


Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=hOGtvAtMSL82
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
Disk 84.44 GB available
+ Code + Text
RAM
Disk
def compareModel(x_train,X_test,y_train,y_test):
decisionTree(x_train,x_test,y_train,y_test)
print('-'*100)
RandomForest(x_train,x_test,y_train,y_test)
print('-'*100)
XGB(x_train,x_test,y_train,y_test)
print('-'*100)
KNN(x_train,x_test,y_train,y_test)
print('-'*100)
[ ] compareModel(x_train,x_test,y_train,y_test)
NameError Traceback (most recent call last)
<ipython-input-22-3eb19b242d41> in <cell line: 1>()
----> 1 compareModel(x_train,x_test,y_train,y_test)
NameError: name 'x_train' is not defined
SEARCH STACK OVERFLOW
[ ] yPred = classifier.predict(x_test)
print(accuracy_score(y_pred,y_test))
print("ANN Model")
print("Confution Matrix")
Type here to search
30°C Partly cloudy
ENG IN 7:28 AM 4/22/2023


Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1mJ2Lt5CzIT5wgs1MDMO844JrzXfaJ6qU#scrollTo=2W0-ePYVWJce
Untitled0.ipynb
File Edit View Insert Runtime Tools Help Last saved at 12:36 AM
Comment
Share
Files
sample_data
+ Code + Text
RAM
Disk
[ ]
----> 1 compareModel(x_train,x_test,y_train,y_test)
NameError: name 'x_train' is not defined
SEARCH STACK OVERFLOW
yPred = classifier.predict(x_test)
print(accuracy_score(y_pred,y_test))
print("ANN Model")
print("Confution_Matrix")
print(confution_matrix(y_test,y_pred))
print("Classification Report")
print(classification_report(y_test,y_pred))
NameError Traceback (most recent call last)
<ipython-input-28-b03b4a9dca55> in <cell line: 1>()
----> 1 yPred = classifier.predict(x_test)
2 print(accuracy_score(y_pred,y_test))
3 print("ANN Model")
4 print("Confution_Matrix")
5 print(confution_matrix(y_test,y_pred))
NameError: name 'classifier' is not defined
SEARCH STACK OVERFLOW
Disk 84.44 GB available
Type here to search
Live
ENG
IN
7:29 AM
4/22/2023