Decision Tree And Random Numbers

```
#Load Libraries
import pandas as pd
import numpy as np
from sklearn.tree import DecisionTreeClassifier
df=pd.read_csv("/content/demodt.txt",sep=",")
features=pd.DataFrame(columns=["Literacy","Cleanliness","Crime_Rate"])
features["Literacy"]=df["Literacy"]
features["Cleanliness"]=df["Cleanliness"]
features["Crime_Rate"]=df["Crime_Rate"]
target=df["Good"]
model=DecisionTreeClassifier()
model.fit(features,target)
Literacy=int(input("Enter the Lietracy Rate : "))
Cleanliness=int(input("Enter the Cleanliness : "))
Crime_Rate=int(input("Enter the Crime Rate : "))
```

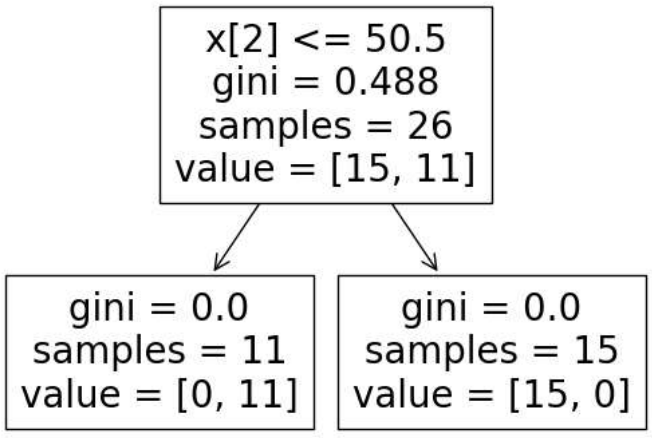
```
Enter the Lietracy Rate : 15
Enter the Cleanliness : 35
Enter the Crime Rate : 24
```

```
p=model.predict([[Literacy,Cleanliness,Crime_Rate]])
if(p==1):
  print("The State is Good ")
else:
  print("The State is Bad ")
```

```
The State is Good
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but DecisionTreeClassifie
  warnings.warn(
```

```
from sklearn.tree import export_graphviz,plot_tree
plot_tree(model)
```

```
[Text(0.5, 0.75, 'x[2] <= 50.5\ngini = 0.488\nsamples = 26\nvalue = [15, 11]'),
 Text(0.25, 0.25, 'gini = 0.0\nsamples = 11\nvalue = [0, 11]'),
 Text(0.75, 0.25, 'gini = 0.0\nsamples = 15\nvalue = [15, 0]')]
```



```
#Load Libraries
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestClassifier
df=pd.read_csv("/content/demodt.txt",sep=",")
features=pd.DataFrame(columns=["Literacy","Cleanliness","Crime_Rate"])
features["Literacy"]=df["Literacy"]
features["Cleanliness"]=df["Cleanliness"]
features["Crime_Rate"]=df["Crime_Rate"]
target=df["Good"]
model=RandomForestClassifier()
model.fit(features,target)
Literacy=int(input("Enter the Lietracy Rate : "))
Cleanliness=int(input("Enter the Cleanliness : "))
Crime_Rate=int(input("Enter the Crime Rate : "))
```

```
Enter the Lietracy Rate : 32
Enter the Cleanliness : 15
Enter the Crime Rate : 41
```

```
p=model.predict([[Literacy,Cleanliness,Crime_Rate]])
if(p==1):
  print("The State is Good ")
else:
  print("The State is Bad ")
```

```
The State is Good
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but RandomForestClassifie
  warnings.warn(
```

```
#Saving all decision trees of RandomForest
import os
output_dir="tree_visualizations"
os.makedirs(output_dir,exist_ok=True)


for i , tree in enumerate(model.estimators_):
  tree_dot_file=os.path.join(output_dir,f"tree_{i}.dot")
  tree_png_file=os.path.join(output_dir,f"tree_{i}.png")


from sklearn.tree import export_graphviz
export_graphviz(tree, out_file=tree_dot_file, feature_names=["Literacy", "Cleanliness", "Crime_Rate"],
                    class_names=[str(cls) for cls in model.classes_],filled=True,rounded=True)


command=f"dot -Tpng {tree_dot_file} -o {tree_png_file}"
os.system(command)
print(f"tree {i} visualization saved to {tree_png_file}")
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
<ipython-input-4-9a824196e64f> in <cell line: 1>()
----> 1 command=f"dot -Tpng {tree_dot_file} -o {tree_png_file}"
      2 os.system(command)
      3 print(f"tree {i} visualization saved to {tree_png_file}")

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

```
#Profit prediction using LinearRegression
import pandas as pd
import numpy as np
from sklearn.linear_model import LinearRegression
model=LinearRegression()
df=pd.read_csv("/content/icecreams.csv")
t=np.array([i for i in df["Temperature"]]).reshape(-1,1)
p=df["Ice Cream Profits"]
model.fit(t,p)
n=int(input("Enter Temperature:"))
profit=(model.predict([[n]]))
print(profit)
if(profit<-5 and profit>-10):
  print("Sales are very low, decrease production")
elif(profit>-5 and profit<10):
  print("Sales are linear please spend some more on marketing")
elif(profit>10):
  print("Sales are rapid, increase production")
else:
  print("Margins are negative, take a break from business :")
```

```
Enter Temperature:22
[-7.47396957]
Sales are very low, decrease production
```

```
import pandas as pd
import seaborn as sns
df=pd.read_csv("/content/dummy_data.csv")
sns.relplot(data=df,x="age",y="time_spent",hue="platform",style="interests")
df
```

	age	gender	time_spent	platform	interests	location	demographics	profession	i
						United		Software	

```
df=pd.read_csv("/content/data.csv",encoding="latin1")
df
```

```
---------------------------------------------------------------------------
FileNotFoundError                         Traceback (most recent call last)
<ipython-input-8-b4ceaf13adba> in <cell line: 1>()
----> 1 df=pd.read_csv("/content/data.csv",encoding="latin1")
      2 df

6 frames
/usr/local/lib/python3.10/dist-packages/pandas/io/common.py in get_handle(path_or_buf,
mode, encoding, compression, memory_map, is_text, errors, storage_options)
    854         if ioargs.encoding and "b" not in ioargs.mode:
    855             # Encoding
--> 856             handle = open(
    857                 handle,
    858                 ioargs.mode,

FileNotFoundError: [Errno 2] No such file or directory: '/content/data.csv'
```

```
df3=df2.loc[df2["location"]=="Hyderabad"]
df3
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
<ipython-input-9-3504b5e9b546> in <cell line: 1>()
----> 1 df3=df2.loc[df2["location"]=="Hyderabad"]
      2 df3

NameError: name 'df2' is not defined
```

```
sns.relplot(data=df3,x="location",y="no2",marker="H",hue="location",color="black",palette="bright")
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
```