In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

In [3]:
```
data=pd.read_csv(r"C:\Users\user\Desktop\vicky\C8_loan-test.csv")
```

In [4]:
```
data.head()
```

Out[4]:

Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_A
Male	Yes	0	Graduate	No	5720	0	110.0	
Male	Yes	1	Graduate	No	3076	1500	126.0	
Male	Yes	2	Graduate	No	5000	1800	208.0	
Male	Yes	2	Graduate	No	2340	2546	100.0	
Male	No	0	Not Graduate	No	3276	0	78.0	

In [5]:

```
data.head
```

Out[5]:

```
<bound method NDFrame.head of       Loan_ID Gender Married Dependents     Education Self_Empl
oyed  \
0    LP001015   Male     Yes          0      Graduate            No   
1    LP001022   Male     Yes          1      Graduate            No   
2    LP001031   Male     Yes          2      Graduate            No   
3    LP001035   Male     Yes          2      Graduate            No   
4    LP001051   Male      No          0  Not Graduate            No   
..        ...    ...     ...        ...           ...           ...   
362  LP002971   Male     Yes         3+  Not Graduate           Yes   
363  LP002975   Male     Yes          0      Graduate            No   
364  LP002980   Male      No          0      Graduate            No   
365  LP002986   Male     Yes          0      Graduate            No   
366  LP002989   Male      No          0      Graduate           Yes   

     ApplicantIncome  CoapplicantIncome  LoanAmount  Loan_Amount_Term  \
0               5720                  0       110.0             360.0   
1               3076               1500       126.0             360.0   
2               5000               1800       208.0             360.0   
3               2340               2546       100.0             360.0   
4               3276                  0        78.0             360.0   
..               ...                ...         ...               ...   
362             4009               1777       113.0             360.0   
363             4158                709       115.0             360.0   
364             3250               1993       126.0             360.0   
365             5000               2393       158.0             360.0   
366             9200                  0        98.0             180.0   

     Credit_History Property_Area  
0               1.0         Urban  
1               1.0         Urban  
2               1.0         Urban  
3               NaN         Urban  
4               1.0         Urban  
..              ...           ...  
362             1.0         Urban  
363             1.0         Urban  
364             NaN     Semiurban  
365             1.0         Rural  
366             1.0         Rural  

[367 rows x 12 columns]>
```

In [17]:

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 367 entries, 0 to 366
Data columns (total 12 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   Loan_ID            367 non-null    object 
 1   Gender             356 non-null    object 
 2   Married            367 non-null    object 
 3   Dependents         357 non-null    object 
 4   Education          367 non-null    object 
 5   Self_Employed      344 non-null    object 
 6   ApplicantIncome    367 non-null    int64  
 7   CoapplicantIncome  367 non-null    int64  
 8   LoanAmount         362 non-null    float64
 9   Loan_Amount_Term   361 non-null    float64
 10  Credit_History     338 non-null    float64
 11  Property_Area      367 non-null    object 
dtypes: float64(3), int64(2), object(7)
memory usage: 34.5+ KB
```

In [36]:

```
data1=data[['ApplicantIncome','CoapplicantIncome','Education']]
```

In [37]:

```
data1['Education'].value_counts()
```

Out[37]:

```
Yes    283
No      84
Name: Education, dtype: int64
```

In [38]:

```
x=data1.drop('Education',axis=1)
y=data1['Education']
```

In [39]:

```
g1={"Education":{"Graduate":0,"Not Graduate":1,}}

data1=data1.replace(g1)
print(data)
```

```
      Loan_ID Gender Married Dependents Education Self_Employed  \
0    LP001015   Male     Yes          0       Yes            No   
1    LP001022   Male     Yes          1       Yes            No   
2    LP001031   Male     Yes          2       Yes            No   
3    LP001035   Male     Yes          2       Yes            No   
4    LP001051   Male      No          0        No            No   
..        ...    ...     ...        ...       ...           ...   
362  LP002971   Male     Yes         3+        No           Yes   
363  LP002975   Male     Yes          0       Yes            No   
364  LP002980   Male      No          0       Yes            No   
365  LP002986   Male     Yes          0       Yes            No   
366  LP002989   Male      No          0       Yes           Yes   

     ApplicantIncome  CoapplicantIncome  LoanAmount  Loan_Amount_Term  \
0               5720                  0       110.0             360.0   
1               3076               1500       126.0             360.0   
2               5000               1800       208.0             360.0   
3               2340               2546       100.0             360.0   
4               3276                  0        78.0             360.0   
..               ...                ...         ...               ...   
362             4009               1777       113.0             360.0   
363             4158                709       115.0             360.0   
364             3250               1993       126.0             360.0   
365             5000               2393       158.0             360.0   
366             9200                  0        98.0             180.0   

     Credit_History Property_Area  
0               1.0         Urban  
1               1.0         Urban  
2               1.0         Urban  
3               NaN         Urban  
4               1.0         Urban  
..              ...           ...  
362             1.0         Urban  
363             1.0         Urban  
364             NaN     Semiurban  
365             1.0         Rural  
366             1.0         Rural  

[367 rows x 12 columns]
```

In [40]:

```
from sklearn.model_selection import train_test_split
```

In [41]:

```
x_train,x_test,y_train,y_test=train_test_split(x,y,train_size=0.70)
```

In [42]:

```
from sklearn.ensemble import RandomForestClassifier
```

In [43]:

```
rfc=RandomForestClassifier()
rfc.fit(x_train,y_train)
```

Out[43]:

```
RandomForestClassifier()
```

In [44]:

```
parameters = {'max_depth':[1,2,3,4,5],
              'min_samples_leaf':[5,10,15,20,25],
              'n_estimators':[10,20,30,40,50]

}
```

In [45]:

```
from sklearn.model_selection import GridSearchCV
grid_search=GridSearchCV(estimator=rfc,param_grid=parameters,cv=2,scoring="accuracy")
grid_search.fit(x_train,y_train)
```

Out[45]:

```
GridSearchCV(cv=2, estimator=RandomForestClassifier(),
             param_grid={'max_depth': [1, 2, 3, 4, 5],
                         'min_samples_leaf': [5, 10, 15, 20, 25],
                         'n_estimators': [10, 20, 30, 40, 50]},
             scoring='accuracy')
```

In [46]:

```
grid_search.best_score_
```

Out[46]:

```
0.796875
```

In [47]:

```
from sklearn.tree import plot_tree
```

In [48]:

```
rfc_best=grid_search.best_estimator_
```

In [49]:

```
plt.figure(figsize=(80,40))
plot_tree(rfc_best.estimators_[5],feature_names=x.columns,class_names=['Yes','No'],filled=True
```

Out[49]:

```
[Text(2418.0, 1902.6000000000001, 'CoapplicantIncome <= 2914.0\ngini = 0.35\nsamples = 163\nv
alue = [58, 198]\nclass = No'),
 Text(1488.0, 1359.0, 'CoapplicantIncome <= 869.0\ngini = 0.384\nsamples = 133\nvalue = [55,
157]\nclass = No'),
 Text(744.0, 815.4000000000001, 'ApplicantIncome <= 5324.5\ngini = 0.295\nsamples = 72\nvalue
= [21, 96]\nclass = No'),
 Text(372.0, 271.79999999999995, 'gini = 0.355\nsamples = 47\nvalue = [18, 60]\nclass = No'),
 Text(1116.0, 271.79999999999995, 'gini = 0.142\nsamples = 25\nvalue = [3, 36]\nclass = No'),
 Text(2232.0, 815.4000000000001, 'CoapplicantIncome <= 2691.5\ngini = 0.46\nsamples = 61\nval
ue = [34, 61]\nclass = No'),
 Text(1860.0, 271.79999999999995, 'gini = 0.442\nsamples = 55\nvalue = [29, 59]\nclass = N
o'),
 Text(2604.0, 271.79999999999995, 'gini = 0.408\nsamples = 6\nvalue = [5, 2]\nclass = Yes'),
 Text(3348.0, 1359.0, 'ApplicantIncome <= 2561.5\ngini = 0.127\nsamples = 30\nvalue = [3, 41]
\nclass = No'),
 Text(2976.0, 815.4000000000001, 'gini = 0.278\nsamples = 8\nvalue = [2, 10]\nclass = No'),
 Text(3720.0, 815.4000000000001, 'CoapplicantIncome <= 3722.0\ngini = 0.061\nsamples = 22\nva
lue = [1, 31]\nclass = No'),
 Text(3348.0, 271.79999999999995, 'gini = 0.219\nsamples = 6\nvalue = [1, 7]\nclass = No'),
 Text(4092.0, 271.79999999999995, 'gini = 0.0\nsamples = 16\nvalue = [0, 24]\nclass = No')]
```

In [ ]:

```
```