In [81]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [219]:
```
df=pd.read_csv(r"C:\USERS\user\Downloads\C8_loan-train - C8_loan-train.csv")
```

Out[219]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	Coap
0	LP001002	Male	No	0	Graduate	No	5849	
1	LP001003	Male	Yes	1	Graduate	No	4583	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	
4	LP001008	Male	No	0	Graduate	No	6000	
...	...	...	...	...	...	...	...	...
609	LP002978	Female	No	0	Graduate	No	2900	
610	LP002979	Male	Yes	3+	Graduate	No	4106	
611	LP002983	Male	Yes	1	Graduate	No	8072	
612	LP002984	Male	Yes	2	Graduate	No	7583	
613	LP002990	Female	No	0	Graduate	Yes	4583	

614 rows × 13 columns

In [220]:
```
df.columns
```

Out[220]:
```
Index(['Loan_ID', 'Gender', 'Married', 'Dependents', 'Education',
       'Self_Employed', 'ApplicantIncome', 'CoapplicantIncome', 'LoanAmount',
       'Loan_Amount_Term', 'Credit_History', 'Property_Area', 'Loan_Status'],
      dtype='object')
```

In [230]:
```
df=df.head(15)
```

Out[230]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	Coapp
0	LP001002	Male	No	0	Graduate	No	5849	
1	LP001003	Male	Yes	1	Graduate	No	4583	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	
3	LP001006	Male	Yes	0	Not Graduate	No	2583	
4	LP001008	Male	No	0	Graduate	No	6000	
5	LP001011	Male	Yes	2	Graduate	Yes	5417	
6	LP001013	Male	Yes	0	Not Graduate	No	2333	
7	LP001014	Male	Yes	3+	Graduate	No	3036	
8	LP001018	Male	Yes	2	Graduate	No	4006	
9	LP001020	Male	Yes	1	Graduate	No	12841	
10	LP001024	Male	Yes	2	Graduate	No	3200	
11	LP001027	Male	Yes	2	Graduate	NaN	2500	
12	LP001028	Male	Yes	2	Graduate	No	3073	
13	LP001029	Male	No	0	Graduate	No	1853	
14	LP001030	Male	Yes	2	Graduate	No	1299	

In [231]:
```
a=df[['Property_Area', 'ApplicantIncome', 'CoapplicantIncome','Loan_Amount_Ter
```

Out[231]:

	Property_Area	ApplicantIncome	CoapplicantIncome	Loan_Amount_Term	Credit_History
0	Urban	5849	0.0	360.0	1.0
1	Rural	4583	1508.0	360.0	1.0
2	Urban	3000	0.0	360.0	1.0
3	Urban	2583	2358.0	360.0	1.0
4	Urban	6000	0.0	360.0	1.0
5	Urban	5417	4196.0	360.0	1.0
6	Urban	2333	1516.0	360.0	1.0
7	Semiurban	3036	2504.0	360.0	0.0
8	Urban	4006	1526.0	360.0	1.0
9	Semiurban	12841	10968.0	360.0	1.0
10	Urban	3200	700.0	360.0	1.0
11	Urban	2500	1840.0	360.0	1.0
12	Urban	3073	8106.0	360.0	1.0
13	Rural	1853	2840.0	360.0	1.0
14	Urban	1299	1086.0	120.0	1.0

In [232]:
```
[illegible]
```

Out[232]:
```
Urban        11
Semiurban     2
Rural         2
Name: Property_Area, dtype: int64
```

In [210]:
```
x=a.drop('Property_Area',axis=1)
```

In [233]:
```
g1={"Property_Area":{'Urban':1,'Semiurban':2,'Rural':5}}
a=a.replace(g1)
```

```
    Property_Area  ApplicantIncome  CoapplicantIncome  Loan_Amount_Term  \
0               1             5849                0.0             360.0
1               5             4583             1508.0             360.0
2               1             3000                0.0             360.0
3               1             2583             2358.0             360.0
4               1             6000                0.0             360.0
5               1             5417             4196.0             360.0
6               1             2333             1516.0             360.0
7               2             3036             2504.0             360.0
8               1             4006             1526.0             360.0
9               2            12841            10968.0             360.0
10              1             3200              700.0             360.0
11              1             2500             1840.0             360.0
12              1             3073             8106.0             360.0
13              5             1853             2840.0             360.0
14              1             1299             1086.0             120.0

    Credit_History
0              1.0
1              1.0
2              1.0
3              1.0
4              1.0
5              1.0
6              1.0
7              0.0
8              1.0
9              1.0
10             1.0
11             1.0
12             1.0
13             1.0
14             1.0
```

In [234]:
```
from sklearn.model_selection import train_test_split
```

In [235]:
```
from sklearn.ensemble import RandomForestClassifier

rfc=RandomForestClassifier()
```

Out[235]:
```
RandomForestClassifier()
```

In [236]:
```
parameters={'max_depth':[1,2,3,4,5],
            'min_samples_leaf':[5,10,15,20,25],
```

In [237]:
```
from sklearn.model_selection import GridSearchCV

grid_search=GridSearchCV(estimator=rfc,param_grid=parameters,cv=2,scoring="acc
```

```
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\model_selection\_split.py:
666: UserWarning: The least populated class in y has only 1 members, which is
less than n_splits=2.
  warnings.warn(("The least populated class in y has only %d"
```

Out[237]:
```
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 [238]:

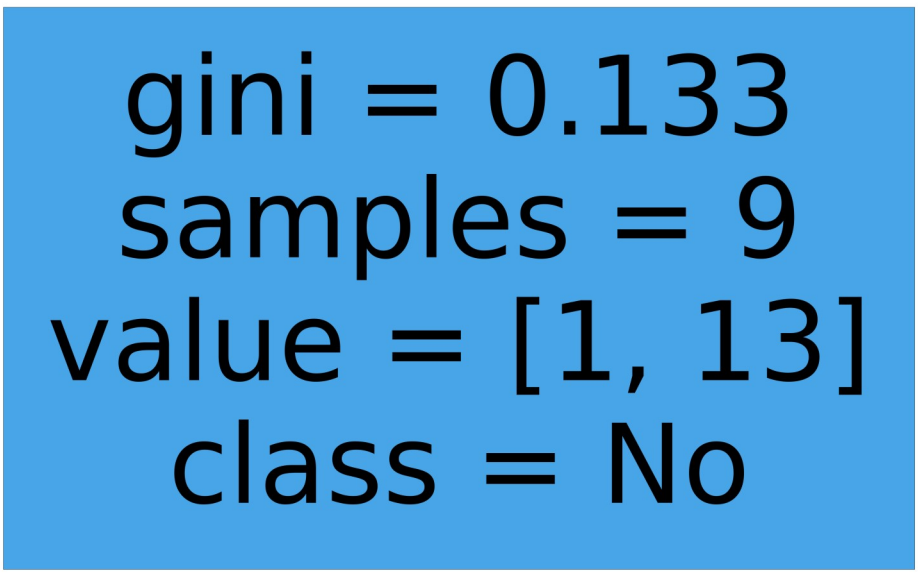
Out[238]:
```
0.9285714285714286
```

In [239]:

In [240]:
```
from sklearn.tree import plot_tree

plt.figure(figsize=(80,40))
```

Out[240]:
```
[Text(2232.0, 1087.2, 'gini = 0.133\nsamples = 9\nvalue = [1, 13]\nclass = No
')]
```



In [ ]: