```
In [1]:
         import numpy as np
         import pandas as pd
         import matplotlib.pyplot as plt
         import seaborn as sea
In [2]: df = pd.read csv(r"C:\Users\user\Downloads\C8 loan-train.csv")
Out[2]:
               Loan_ID Gender Married Dependents Education Self_Employed ApplicantIncome CoapplicantIncome LoanAmount Loan_Amou
            0 LP001002
                                                                                     5849
                          Male
                                   No
                                                    Graduate
                                                                       No
                                                                                                        0.0
                                                                                                                   NaN
            1 LP001003
                                                    Graduate
                                                                                     4583
                                                                                                     1508.0
                                                                                                                  128.0
                          Male
                                   Yes
                                                1
                                                                       No
            2 LP001005
                          Male
                                                0
                                                    Graduate
                                                                      Yes
                                                                                     3000
                                                                                                        0.0
                                                                                                                   66.0
                                   Yes
                                                        Not
            3 LP001006
                                                                                     2583
                                                                                                     2358.0
                                                                                                                  120.0
                          Male
                                   Yes
                                                0
                                                                       No
                                                    Graduate
            4 LP001008
                                                0
                                                    Graduate
                                                                                     6000
                                                                                                        0.0
                                                                                                                  141.0
                          Male
                                   No
                                                                       No
                                    ...
                                                                                                        ...
                                                                                                                     ...
          609 LP002978
                                                0
                                                    Graduate
                                                                                     2900
                                                                                                        0.0
                                                                                                                   71.0
                        Female
                                   No
                                                                       No
          610 LP002979
                          Male
                                               3+
                                                    Graduate
                                                                       No
                                                                                     4106
                                                                                                        0.0
                                                                                                                   40.0
                                   Yes
          611 LP002983
                          Male
                                   Yes
                                                    Graduate
                                                                       No
                                                                                     8072
                                                                                                      240.0
                                                                                                                  253.0
                                                2
                                                                                                        0.0
                                                                                                                  187.0
          612 LP002984
                          Male
                                                    Graduate
                                                                                     7583
                                   Yes
                                                                       No
          613 LP002990
                       Female
                                   No
                                                    Graduate
                                                                      Yes
                                                                                     4583
                                                                                                        0.0
                                                                                                                  133.0
         614 rows × 13 columns
In [3]: df.info()
         <class 'pandas.core.frame.DataFrame'>
         RangeIndex: 614 entries, 0 to 613
         Data columns (total 13 columns):
                                   Non-Null Count Dtype
          # Column
          0
              Loan_ID
                                                     object
                                   614 non-null
          1
              Gender
                                   601 non-null
                                                     object
              Married
                                   611 non-null
                                                     object
          2
          3
              Dependents
                                   599 non-null
                                                     object
                                   614 non-null
          4
              Education
                                                     object
              Self Employed
                                   582 non-null
                                                     object
              ApplicantIncome
                                   614 non-null
                                                     int64
                                   614 non-null
                                                     float64
          7
              CoapplicantIncome
          8
              LoanAmount
                                   592 non-null
                                                     float64
              Loan_Amount_Term
                                   600 non-null
                                                     float64
              Credit_History
                                   564 non-null
                                                     float64
          10
          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
```

In [4]: df2 = df.fillna(0)

In [5]: df2

Out[5]:

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amou
0	LP001002	Male	No	0	Graduate	No	5849	0.0	0.0	
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										

In [6]: df2.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	614 non-null	object			
2	Married	614 non-null	object			
3	Dependents	614 non-null	object			
4	Education	614 non-null	object			
5	Self_Employed	614 non-null	object			
6	ApplicantIncome	614 non-null	int64			
7	CoapplicantIncome	614 non-null	float64			
8	LoanAmount	614 non-null	float64			
9	Loan_Amount_Term	614 non-null	float64			
10	Credit_History	614 non-null	float64			
11	Property_Area	614 non-null	object			
12	Loan_Status	614 non-null	object			
<pre>dtypes: float64(4), int64(1), object(8)</pre>						

In [7]: df3 = df2[['ApplicantIncome','CoapplicantIncome','LoanAmount','Loan_Amount_Term','Credit_History','Loan_Status
df3

Out[7]:

	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Loan_Status
0	5849	0.0	0.0	360.0	1.0	Υ
1	4583	1508.0	128.0	360.0	1.0	N
2	3000	0.0	66.0	360.0	1.0	Υ
3	2583	2358.0	120.0	360.0	1.0	Υ
4	6000	0.0	141.0	360.0	1.0	Υ
609	2900	0.0	71.0	360.0	1.0	Υ
610	4106	0.0	40.0	180.0	1.0	Υ
611	8072	240.0	253.0	360.0	1.0	Υ
612	7583	0.0	187.0	360.0	1.0	Υ
613	4583	0.0	133.0	360.0	0.0	N

614 rows × 6 columns

memory usage: 62.5+ KB

In [8]: df3.info()

```
<class 'pandas.core.frame.DataFrame'>
          RangeIndex: 614 entries, 0 to 613
          Data columns (total 6 columns):
                                   Non-Null Count Dtype
           #
              Column
           0
               ApplicantIncome
                                   614 non-null
                                                     int64
               CoapplicantIncome
                                                     float64
           1
                                   614 non-null
               LoanAmount
                                   614 non-null
                                                     float64
               Loan_Amount_Term
                                   614 non-null
                                                     float64
               Credit_History
                                   614 non-null
                                                     float64
               Loan Status
                                   614 non-null
                                                     object
          dtypes: float64(4), int64(1), object(1)
          memory usage: 28.9+ KB
In [10]: df3['Loan_Status'].value_counts()
Out[10]: Y
               422
               192
          Name: Loan_Status, dtype: int64
In [11]: | ch = {"Loan_Status":{'Y':1,'N':0}}
          df3 = df3.replace(ch)
         df3
Out[11]:
               ApplicantIncome
                              CoapplicantIncome LoanAmount Loan Amount Term Credit History Loan Status
            0
                                                                                                    1
                                           0.0
                                                                       360.0
                                         1508.0
                                                                                                    0
             1
                         4583
                                                      128.0
                                                                       360.0
                                                                                       1.0
            2
                         3000
                                           0.0
                                                       66.0
                                                                       360.0
                                                                                       1.0
             3
                         2583
                                         2358.0
                                                      120.0
                                                                        360.0
                                                                                       1.0
                         6000
                                                                       360.0
             4
                                           0.0
                                                      141.0
                                                                                       1.0
           609
                         2900
                                           0.0
                                                       71.0
                                                                        360.0
                                                                                       1.0
           610
                         4106
                                           0.0
                                                       40.0
                                                                        180.0
                                                                                       1.0
           611
                         8072
                                         240.0
                                                      253.0
                                                                       360.0
                                                                                       1.0
           612
                         7583
                                           0.0
                                                      187.0
                                                                        360.0
                                                                                       1.0
           613
                         4583
                                           0.0
                                                      133.0
                                                                        360.0
                                                                                       0.0
                                                                                                    n
          614 rows × 6 columns
In [14]: | x = df3[['ApplicantIncome','CoapplicantIncome','LoanAmount','Loan_Amount_Term','Credit_History']]
          y = df3['Loan_Status']
In [15]: from sklearn.model_selection import train_test_split
In [16]: x_train,x_test,y_train,y_test = train_test_split(x,y,train_size=0.70)
In [17]: from sklearn.ensemble import RandomForestClassifier
In [18]: rfc = RandomForestClassifier()
          rfc.fit(x_train,y_train)
Out[18]: RandomForestClassifier()
In [19]: parameters = {
              'max_depth':[1,2,3,4,5],
              'min_samples_leaf':[5,10,15,20,25],
              'n_estimators':[10,20,30,40,50]
          }
```

```
In [20]: from sklearn.model_selection import GridSearchCV
In [21]: grid_search = GridSearchCV(estimator=rfc,param_grid=parameters,cv=2,scoring='accuracy')
         grid_search.fit(x_train,y_train)
Out[21]: 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 [22]: grid search.best score
Out[22]: 0.7809280591175831
In [23]: from sklearn.tree import plot_tree
In [24]: rfc best= grid search.best estimator
In [25]: plt.figure(figsize=(80,40))
         plot tree(rfc best.estimators [5],feature names=x.columns,class names=["Yes","No"],filled=True)
Out[25]: [Text(2232.0, 1812.0, 'ApplicantIncome <= 2659.5\ngini = 0.428\nsamples = 280\nvalue = [133, 296]\nclass = N</pre>
         o'),
          Text(1116.0, 1087.2, 'CoapplicantIncome <= 2018.5\ngini = 0.492\nsamples = 68\nvalue = [47, 61]\nclass = N
         ο'),
          Text(1674.0, 362.399999999986, 'gini = 0.329\nsamples = 36\nvalue = [11, 42]\nclass = No'),
          Text(3348.0, 1087.2, 'CoapplicantIncome <= 94.5\ngini = 0.392\nsamples = 212\nvalue = [86, 235]\nclass = N
         o'),
          Text(2790.0, 362.399999999986, 'gini = 0.432\nsamples = 111\nvalue = [54, 117]\nclass = No'),
          Text(3906.0, 362.3999999999986, 'gini = 0.336\nsamples = 101\nvalue = [32, 118]\nclass = No')]
                                                   ApplicantIncome <= 2659.5
                                                          gini = 0.428
                                                         samples = 280
                                                       value = [133, 296]
                                                          class = No
                       CoapplicantIncome <= 2018.5
                                                                              CoapplicantIncome <= 94.5
                               gini = 0.492
                                                                                    gini = 0.392
                              samples = 68
                                                                                   samples = 212
                             value = [47, 61]
                                                                                   value = [86, 235]
                               class = No
                                                                                     class = No
                                                                       gini = 0.432
                                                                                                  gini = 0.336
                 aini = 0.452
                                            aini = 0.329
                 samples = 32
                                            samples = 36
                                                                      samples = 111
                                                                                                 samples = 101
                value = [36, 19]
                                           value = [11, 42]
                                                                     value = [54, 117]
                                                                                                value = [32, 118]
                                                                                                   class = No
                  class = Yes
                                             class = No
                                                                        class = No
In [ ]:
```