In [1]:

import pandas as pd

In [2]:

data = pd.read\_excel("Bank\_Personal\_Loan\_Modelling.xlsx",sheet\_name = 1)

In [3]:

data.head(1) Out[3]:

**ID Age Experience Income ZIP**

**Code**

**Family CCAvg Education Mortgage Personal**

**Loan**

**Securities Account**

**CD**

**Account**

**Online CreditC**

**0** 1 25 1 49 91107 4 1.6 1 0 0 1 0 0

In [4]:

data.isnull().sum() Out[4]:

ID 0

Age 0

Experience 0

Income 0

ZIP Code 0

Family 0

CCAvg 0

Education 0

Mortgage 0

Personal Loan 0

Securities Account 0

CD Account 0

Online 0

CreditCard 0

dtype: int64

In [5]:

df = pd.DataFrame(data) df.CCAvg = df.CCAvg.astype(int)

In [6]:

data.head(1) Out[6]:

**ID Age Experience Income ZIP**

**Code**

**Family CCAvg Education Mortgage Personal**

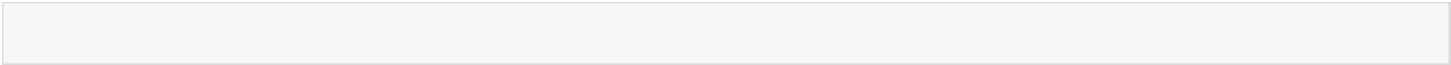
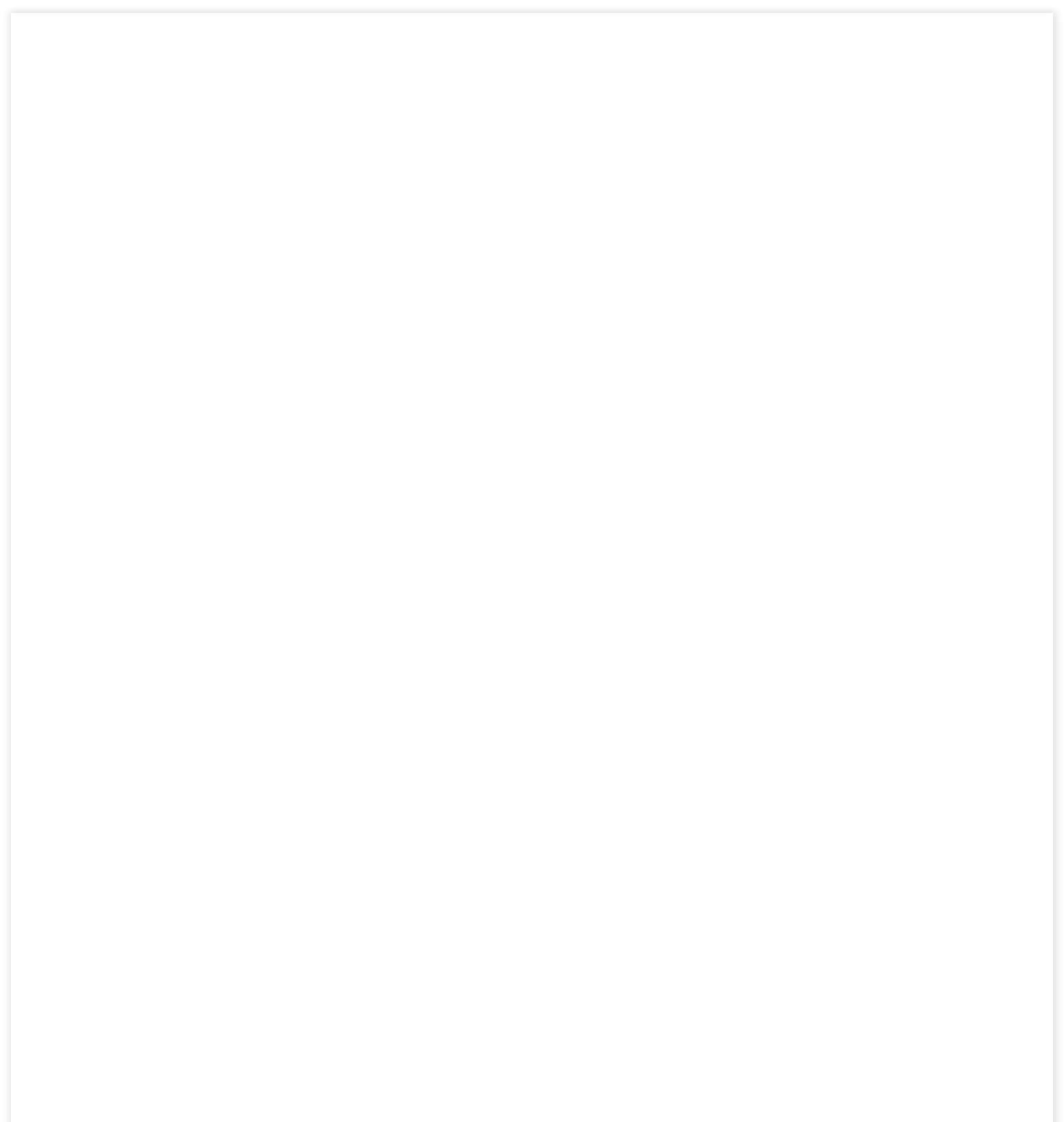
**Loan**

**Securities Account**

**CD**

**Account**

**Online CreditC**



**0** 1 25 1 49 91107 4 1 1 0 0 1 0 0

In [7]:

**from sklearn.ensemble import** RandomForestClassifier

In [8]:

rf\_model=RandomForestClassifier(n\_estimators=1000,max\_features=2,oob\_score=**True**)

In [9]:

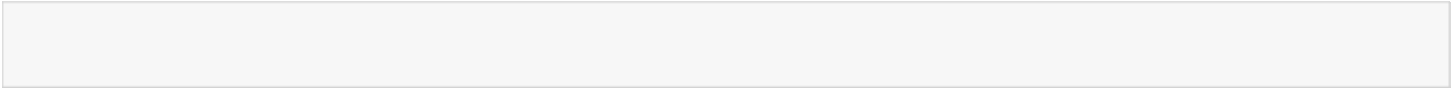


data.columns

Out[9]:

Index(['ID', 'Age', 'Experience', 'Income', 'ZIP Code', 'Family', 'CCAvg', 'Education', 'Mortgage', 'Personal Loan', 'Securities Account', 'CD Account', 'Online', 'CreditCard'],

dtype='object') In [10]:



features=['Age', 'Experience', 'Income', 'Family', 'CCAvg', 'Education', 'Mortgage', 'Securities Account',

'CD Account', 'Online', 'CreditCard']

In [11]:



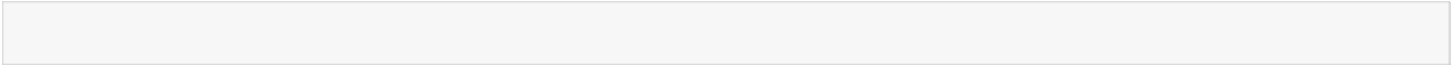
rf\_model.fit(X=data[features],y=data["Personal Loan"])

Out[11]:

RandomForestClassifier(bootstrap=True, ccp\_alpha=0.0, class\_weight=None,

criterion='gini', max\_depth=None, max\_features=2, max\_leaf\_nodes=None, max\_samples=None, min\_impurity\_decrease=0.0, min\_impurity\_split=None, min\_samples\_leaf=1, min\_samples\_split=2, min\_weight\_fraction\_leaf=0.0, n\_estimators=1000, n\_jobs=None, oob\_score=True, random\_state=None, verbose=0, warm\_start=False)

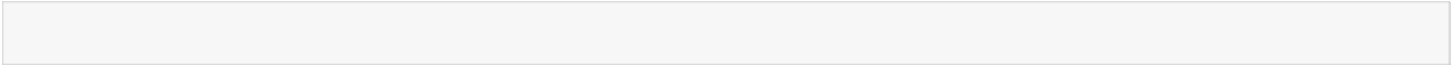
In [12]:



print('OOB Accuracy') print(rf\_model.oob\_score\_);

OOB Accuracy 0.9878

In [13]:



**for** feature,imp **in** zip(features,rf\_model.feature\_importances\_): print(feature,imp)

Age 0.0496890852363932

Experience 0.04932337665613696

Income 0.3574458086321771

Family 0.09762752910743437

CCAvg 0.15403779035754883

Education 0.1628131958606067

Mortgage 0.04706841871088275

Securities Account 0.005912003031798994

CD Account 0.05607504897503181

Online 0.009181081924913606

CreditCard 0.010826661507075655

Income,CCAvg,Education are most significant attributes

In [14]:



**from sklearn import** tree,preprocessing

In [15]:



data.columns

Out[15]:

Index(['ID', 'Age', 'Experience', 'Income', 'ZIP Code', 'Family', 'CCAvg', 'Education', 'Mortgage', 'Personal Loan', 'Securities Account', 'CD Account', 'Online', 'CreditCard'],

dtype='object')

In [16]:

predictors=pd.DataFrame([data['Income'], data['CCAvg'],data['Education']]).T

In [17]:

tree\_model=tree.DecisionTreeClassifier(max\_depth=6)

In [18]:

tree\_model.fit(X=predictors, y=data['Personal Loan']) Out[18]:

DecisionTreeClassifier(ccp\_alpha=0.0, class\_weight=None, criterion='gini',

max\_depth=6, max\_features=None, max\_leaf\_nodes=None, min\_impurity\_decrease=0.0, min\_impurity\_split=None, min\_samples\_leaf=1, min\_samples\_split=2, min\_weight\_fraction\_leaf=0.0, presort='deprecated', random\_state=None, splitter='best')

In [19]:

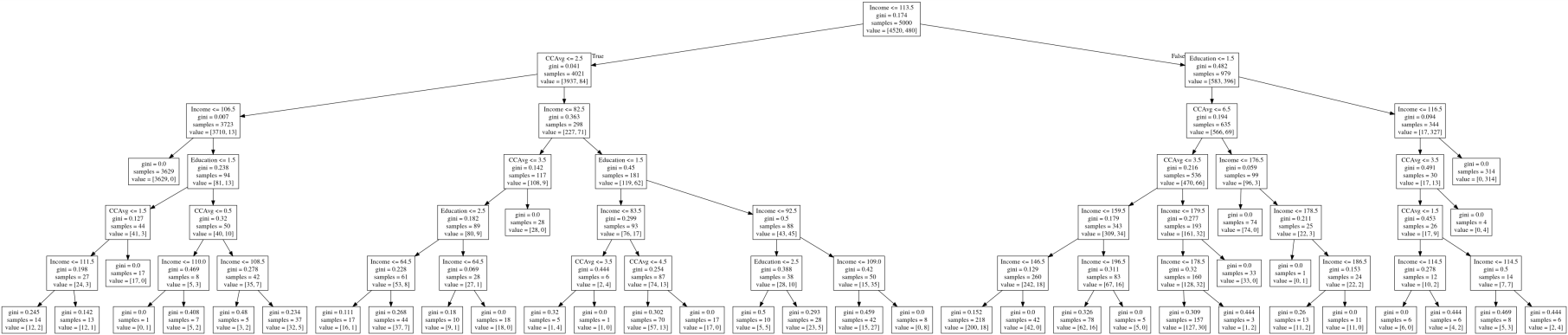
**with** open('bank\_tree.dot','w')**as** f: f=tree.export\_graphviz(tree\_model,feature\_names=['Income','CCAvg','Education'],out\_f

ile=f);

In [20]:

tree\_model.score(X=predictors,y=data['Personal Loan']) Out[20]:

0.9732



**RULES :**

• If income<=113.5 and CCAvg<=2.5 and income<=106.5 then loan sanctioned.

• If income<=106.5 and education<=1.5 and CCavg<=1.5 then loan sanctioned.

• If income>111.5 and CCAvg<=1.5 and Education<=1.5 then loan sanctioned.

• If income<=110.5 and CCAvg<=0.5 and Education<=1.5 then loan not sanctioned.

• If income>110.5 and CCAvg<=0.5 and Education<=1.5 then loan sanctioned.

• If income>108.5 and CCAvg<=0.5 and Education<=1.5 then loan sanctioned.

• If income>116.5 and Education>=1.5 then loan not sanctioned.

• If income<=116.5 and CCAvg>=3.5 and Education<=1.5 then loan sanctioned.

• If income<=114.5 and CCAvg<=1.5 and Education<=1.5 then loan sanctioned.

• If income<=114.5 and CCAvg>=1.5 and Education<=1.5 then loan sanctioned.