

6. Program to implement decision trees using any standard dataset available in the public domain and find the accuracy of the algorithm.

```
import numpy as np
import pandas as pd
from sklearn import tree
```

```
file='loan1.csv'
df=pd.read_csv(file)
```

```
df.head()
```

	Home Owner	Marital Status	Annual Income	Defaulted
	Borrower			
0	Yes	Single 125	No	
1	No	Married 100	No	
2	No	Single 70	No	
3	Yes	Married 120	No	
4	No	Divorced 95	Yes	

```
df.shape
```

```
(10, 4)
```

```
df.info
```

```
<bound method DataFrame.info of Home Owner Marital Status Annual  
Income Defaulted Borrower
```

0	Yes	Single	125	No
1	No	Married	100	No
2	No	Single	70	No
3	Yes	Married	120	No
4	No	Divorced	95	Yes
5	No	Married	60	No
6	Yes	Divorced	220	No
7	No	Single	85	Yes
8	No	Married	75	No
9	No	Single	90	Yes>

```

d={'Yes':1,'No':0}
df['Home Owner']=df['Home Owner'].map(d)
df['Defaulted Borrower']=df['Defaulted Borrower'].map(d)
d1={'Single':0,'Married':1,'Divorced':2}
df['Marital Status']=df['Marital Status'].map(d1)
df.head()

```

	Home Owner	Marital Status	Annual Income	Defaulted Borrower
0	1	0	125	0
1	0	1	100	0
2	0	0	70	0
3	1	1	120	0
4	0	2	95	1

df.columns

```

Index(['Home Owner', 'Marital Status', 'Annual Income', 'Defaulted
Borrower'], dtype='object')

```

```

X=df[['Home Owner', 'Marital Status', 'Annual Income']]
y=df['Defaulted Borrower']
features=list(df.columns[:3])
print(features)

```

```

['Home Owner', 'Marital Status', 'Annual Income']

```

```

model=tree.DecisionTreeClassifier(criterion='entropy',splitter='random
')
model=model.fit(X,y)

```

```

from IPython.display import Image
from sklearn.externals.six import StringIO
import pydotplus

```

```

dot_data=StringIO()

```

```

tree.export_graphviz(model,out_file=dot_data,feature_names=feature
s,filled=True,rounded=True)

```

