19113077 Eshwara.P CSE - 5B Decision Tree

```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split
from sklearn import metrics
```

```
from google.colab import files
dataset = files.upload()
```

Choose Files diabetes.csv

- **diabetes.csv**(application/vnd.ms-excel) - 23873 bytes, last modified: 9/19/2019 - 100% done

```
Saving diabetes.csv to diabetes.csv
```

```
pima = pd.read_csv("diabetes.csv")
```

```
pima.head()
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigre
0	6	148	72	35	0	33.6	
1	1	85	66	29	0	26.6	
2	8	183	64	0	0	23.3	
3	1	89	66	23	94	28.1	
4	0	137	40	35	168	43.1	

```
feature_cols = ['Pregnancies', 'Glucose', 'BloodPressure','SkinThickness','Insulin','BMI',
X = pima[feature_cols]
y = pima.Outcome
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
```

```
clf = DecisionTreeClassifier()

clf = clf.fit(X_train,y_train)

y_pred = clf.predict(X_test)
```

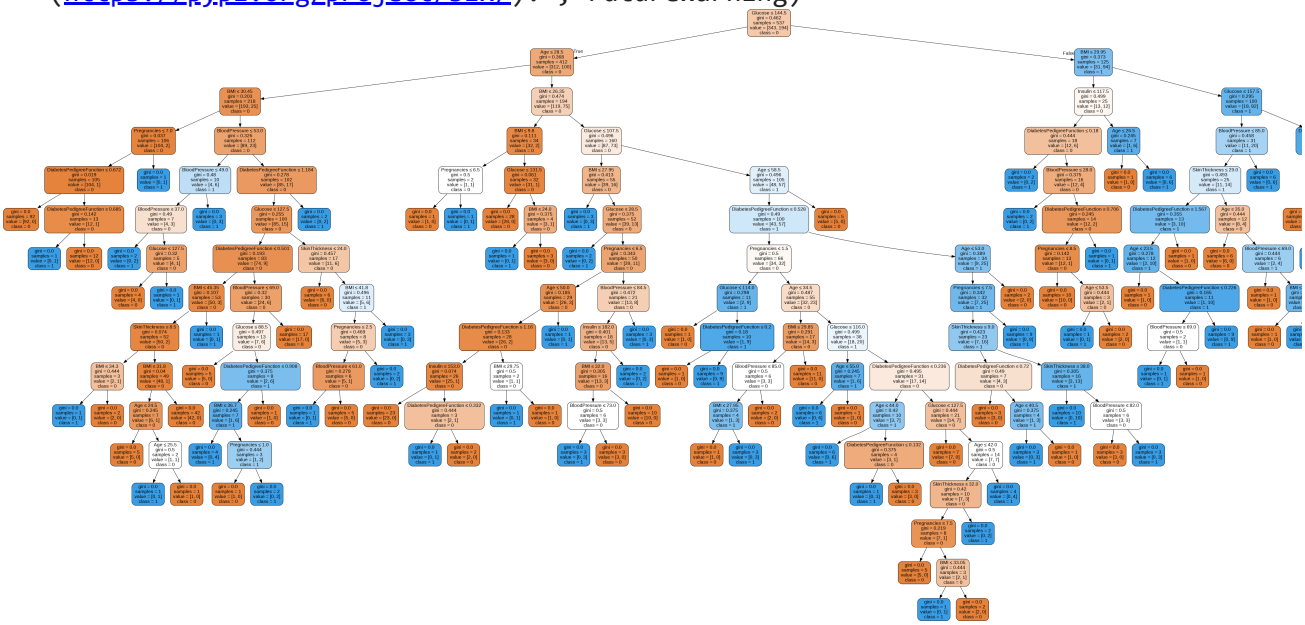
```
print("Accuracy:",metrics.accuracy_score(y_test, y_pred))
```

```
Accuracy: 0.7792207792207793
```

```
from sklearn.externals.six import StringIO
from IPython.display import Image
from sklearn.tree import export_graphviz
```

```
import pydotplus
dot_data = StringIO()
export_graphviz(clf, out_file=dot_data,
                filled=True, rounded=True,
                special_characters=True, feature_names = feature_cols,class_names=['0','1'
graph = pydotplus.graph_from_dot_data(dot_data.getvalue())
graph.write_png('diabetes.png')
Image(graph.create_png())
```

```
/usr/local/lib/python3.7/dist-packages/sklearn/externals/six.py:31: FutureWarning: Th
  "(https://pypi.org/project/six/).", FutureWarning)
```



✓ 0s completed at 4:00 PM