**TASK 1:**

Train the decision tree using the dataset provided in the lab also given below.

**CODE:**

import pandas as pd

import seaborn as sns

from sklearn.tree import DecisionTreeClassifier

from sklearn.model\_selection import train\_test\_split

from sklearn import metrics

from sklearn.metrics import confusion\_matrix

from sklearn.externals.six import StringIO

from IPython.display import Image

from sklearn.tree import export\_graphviz

import pydotplus

pima =  pd.read\_csv("../input/diabetes.csv")

pima.head()

corr = pima.corr()

ax = sns.heatmap(

    corr,

    vmin=-1, vmax=1, center=0,

    cmap=sns.diverging\_palette(20, 220, n=200),

    square=True

)

ax.set\_xticklabels(

    ax.get\_xticklabels(),

    rotation=45,

    horizontalalignment='right'

)

feature\_cols = ['Pregenant', 'Insulin', 'BMI', 'Age', 'Glucose', 'BP', 'Pedigree','skin','label']

x = pima[feature\_cols]

y = pima.Outcome

(X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(x, y, test\_size = 0.3, random\_state=1)

classifier = DecisionTreeClassifier()

classifier = classifier.fit(X\_train, Y\_train)

y\_pred = classifier.predict(X\_test)

print(y\_pred)

confusion\_matrix(Y\_test, y\_pred)

print(confusion\_matrix(Y\_test, y\_pred))

# accuracy

print("Accuracy:", metrics.accuracy\_score(Y\_test,y\_pred))

dot\_data = StringIO()

export\_graphviz(classifier, out\_file=dot\_data,

                filled=True, rounded=True,

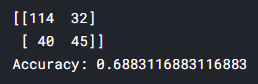
                special\_characters=True, feature\_names = feature\_cols,class\_names=['0','1'])

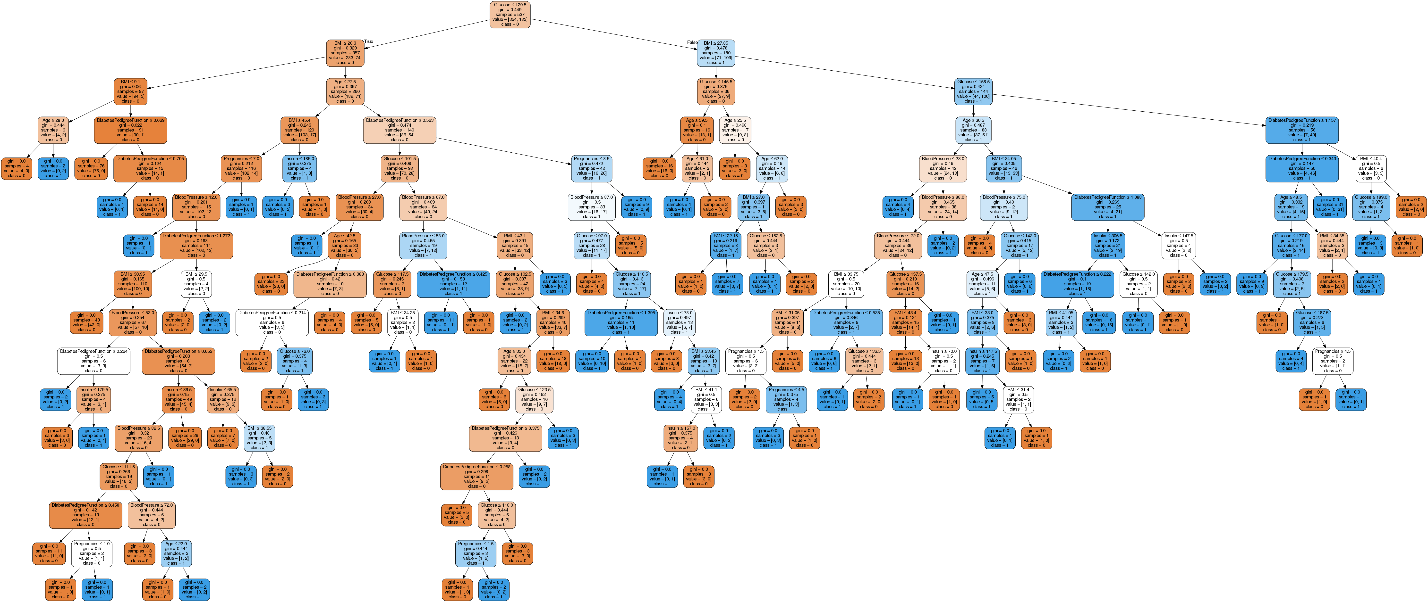
graph = pydotplus.graph\_from\_dot\_data(dot\_data.getvalue())

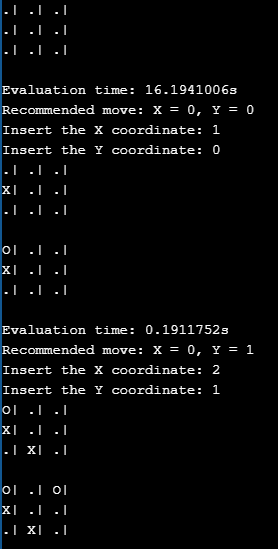
graph.write\_png('diabetes.png')

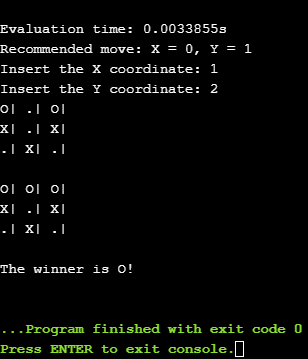
Image(graph.create\_png())

**OUTPUT:**



****

****

****