In [3]:

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

In [4]:

```
col_names = ['pregnant', 'glucose', 'bp', 'skin', 'insulin', 'bmi', 'pedigree', 'age',
'label']

df = pd.read_csv(r"C:\Users\elcot\Downloads\archive (10)\diabetes.csv", header=None, na
mes=col_names)
```

In [6]:

```
df.head()
```

Out[6]:

	pregnant	glucose	bp	skin	insulin	bmi	pedigree	age	label
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

In [7]:

```
#splitting dataset in features and target variable
feature_cols = ['pregnant', 'insulin', 'bmi', 'age','skin','glucose','bp','pedigree']
X = df[feature_cols] # Features
Y = df.label # Target variable
```

In [8]:

```
# Split dataset into training set and test set - 70% training and 30% test
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.3, random_state=1
)
```

In [9]:

```
clf = DecisionTreeClassifier(criterion="entropy", max_depth=3)# Create Decision Tree cl
assifer
```

In [10]:

```
clf = clf.fit(X_train,Y_train)# Train
```

In [11]:

```
#Predict the response for test dataset
Y_pred = clf.predict(X_test)
```

In [12]:

```
print(Y_pred)
```

```
[0 0 0 0 0 0 0 0 0 0 1 0 1 1 0 1 0 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1 0
 0 0 0 0 0 0 1 0 0 1 1 0 0 0 0 1 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 1 1 1 1 0 0
 1 0 1 1 0 1 1 0 0 1 0 1 1 0 1 0 0 0 0 0 1 1 0 0 1 0 0 0 1 1 0 0 1 1 0 0 0
 0 1 0 0 0 0 1 0 1 0 1 0 0 0 0 0 0 1 0 1 0 1 1 0 0 0 1 0 0 1 0 0 1 1 0 0 0
 0 0 0 1 0 0 1 1 0 0 0 0 1 1 0 0 0 0 0 1 1 1 1 1 0 0 1 0 0 0 0 0 0 1 0 1 0
 0 0 1 1 0 0 0 0 0 1 0 0 0 0 1 1 1 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 1 0 0
 0 1 0 1 0 0 0 1 0]
```

In [13]:

```
# Model Accuracy
print("Accuracy:",metrics.accuracy_score(Y_test, Y_pred))
```

```
Accuracy: 0.7705627705627706
```

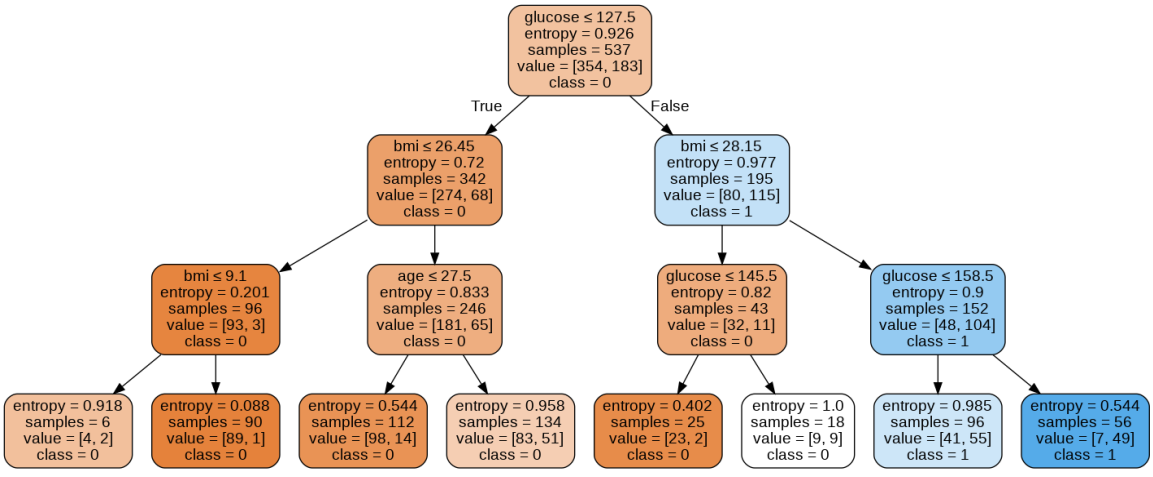
pip install graphviz

pip install pydotplus

In [14]:

```
from 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())
```

Out[14]:



In [14]: