Name - ARYAN BAJAJ

Task 6 - Prediction using Decision Tree Algorithm (Level - Intermediate)

In [1]:
```
# Importing libraries in Python
import sklearn.datasets as datasets
import pandas as pd
```

In [2]:
```
# Loading the iris dataset
iris=datasets.load_iris()
```

In [3]:
```
# Forming the iris dataframe
df=pd.DataFrame(iris.data, columns=iris.feature_names)
print(df.head(5))
```

```
   sepal length (cm)  sepal width (cm)  petal length (cm)  petal width (cm)
0                5.1               3.5                1.4               0.2
1                4.9               3.0                1.4               0.2
2                4.7               3.2                1.3               0.2
3                4.6               3.1                1.5               0.2
4                5.0               3.6                1.4               0.2
```

In [4]:
```
y=iris.target
print(y)
```

```
[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 2 2]
```

Now let us define the Decision Tree Algorithm

In [5]:
```
# Defining the decision tree algorithm
from sklearn.tree import DecisionTreeClassifier
dtree=DecisionTreeClassifier()
dtree.fit(df,y)

print('Decision Tree Classifer Created')
```

```
Decision Tree Classifer Created
```

Let us visualize the Decision Tree to understand it better.

In [6]:
```
# Install required libraries
!pip install pydotplus
!apt-get install graphviz -y
```

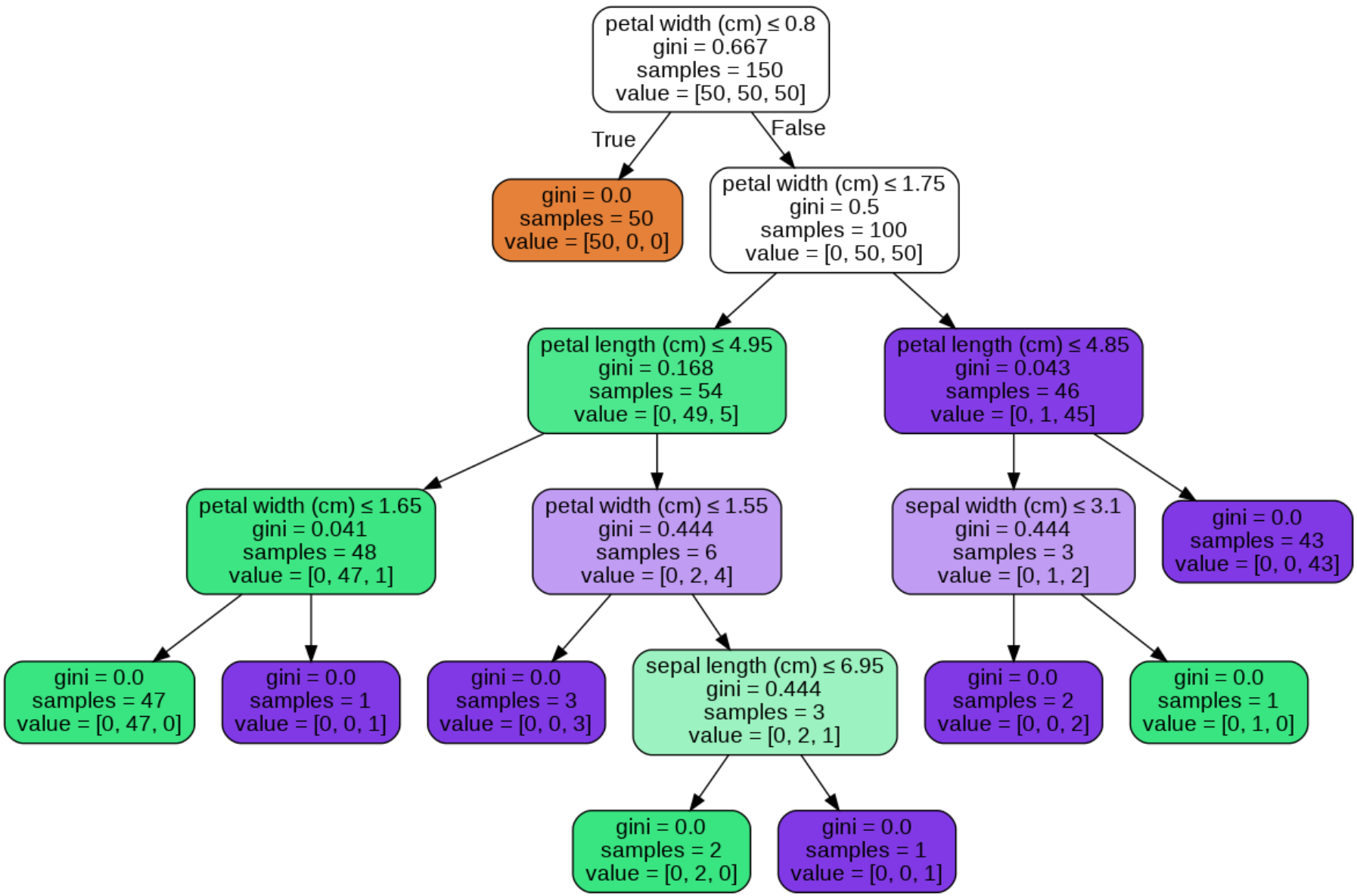
```
Requirement already satisfied: pydotplus in e:\anaconda\lib\site-packages (2.0.2)
Requirement already satisfied: pyparsing>=2.0.1 in e:\anaconda\lib\site-packages (from pydotplus) (2.4.7)

WARNING: Ignoring invalid distribution -tatsmodels (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -ip (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -tatsmodels (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -ip (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -tatsmodels (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -ip (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -tatsmodels (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -ip (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -tatsmodels (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -ip (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -tatsmodels (e:\anaconda\lib\site-packages)
WARNING: Ignoring invalid distribution -ip (e:\anaconda\lib\site-packages)
'apt-get' is not recognized as an internal or external command,
operable program or batch file.
```

In [0]:
```
# Import necessary libraries for graph viz
from sklearn.externals.six import StringIO
from IPython.display import Image
from sklearn.tree import export_graphviz
import pydotplus

# Visualize the graph
dot_data = StringIO()
export_graphviz(dtree, out_file=dot_data, feature_names=iris.feature_names,
                filled=True, rounded=True,
                special_characters=True)
graph = pydotplus.graph_from_dot_data(dot_data.getvalue())
Image(graph.create_png())
```

Out[19]:



You can now feed any new/test data to this classifer and it would be able to predict the right class accordingly.