In [1]:
```
%matplotlib inline     #https://stackoverflow.com/questions/43027980/purpose-of-matplotlib-inline
from sklearn import datasets
from sklearn import tree
import pandas as pd
import numpy as np
import seaborn as sns
import random as rnd
```

In [2]:
```
iris=datasets.load_iris()
df=pd.DataFrame(data=np.c_[iris['data'],iris['target']],columns=iris['feature_names']+['target'])

df.head()
```

Out[2]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
0	5.1	3.5	1.4	0.2	0.0
1	4.9	3.0	1.4	0.2	0.0
2	4.7	3.2	1.3	0.2	0.0
3	4.6	3.1	1.5	0.2	0.0
4	5.0	3.6	1.4	0.2	0.0

In [5]:
```
df.tail()
```

Out[5]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
145	6.7	3.0	5.2	2.3	2.0
146	6.3	2.5	5.0	1.9	2.0
147	6.5	3.0	5.2	2.0	2.0
148	6.2	3.4	5.4	2.3	2.0
149	5.9	3.0	5.1	1.8	2.0

In [6]:
```
sns.set_style("whitegrid") #to set my background of graph to be white
sns.pairplot(df,hue='target')  #gievs the corealtion plot
```

```
C:\Users\vkondala\AppData\Local\Continuum\anaconda3\lib\site-packages\statsmo
dels\nonparametric\kde.py:487: RuntimeWarning: invalid value encountered in t
rue_divide
  binned = fast_linbin(X, a, b, gridsize) / (delta * nobs)
C:\Users\vkondala\AppData\Local\Continuum\anaconda3\lib\site-packages\statsmo
dels\nonparametric\kdetools.py:34: RuntimeWarning: invalid value encountered
in double_scalars
  FAC1 = 2*(np.pi*bw/RANGE)**2
```

Out[6]:
```
<seaborn.axisgrid.PairGrid at 0x1c32b37bcc0>
```

In [15]:
```
x=iris.data[0:150,:]
x.shape
```

Out[15]:
```
(150, 4)
```

In [14]:
```
y=iris.target[0:150]
y.shape
```

Out[14]:
```
(150,)
```

In [16]:
```
setosa_index=rnd.randrange(0,49)    #picks a random number between 0 and 49
test_setosa=[iris.data[setosa_index,:]]
x=np.delete(x,setosa_index,0)
y=np.delete(y,setosa_index,0)

test_setosa,iris.target_names[iris.target[setosa_index]], x.shape,y.shape
```

Out[16]:
```
([array([4.4, 3.2, 1.3, 0.2])], 'setosa', (149, 4), (149,))
```

In [17]:
```
vertosa_index=rnd.randrange(50,99)  #picks a random number between 50 and 09
test_vertosa=[iris.data[vertosa_index,:]]
x=np.delete(x,vertosa_index,0)
y=np.delete(y,vertosa_index,0)

test_vertosa,iris.target_names[iris.target[vertosa_index]], x.shape,y.shape
```

Out[17]:
```
([array([6.3, 2.5, 4.9, 1.5])], 'versicolor', (148, 4), (148,))
```

In [18]:
```
verginica_index=rnd.randrange(100,150)  #picks a random number between 100 and
150
test_vergnica=[iris.data[verginica_index,:]]
x=np.delete(x,verginica_index,0)
y=np.delete(y,verginica_index,0)

test_vergnica,iris.target_names[iris.target[verginica_index]], x.shape,y.shape
```

Out[18]:
```
([array([6.3, 2.8, 5.1, 1.5])], 'virginica', (147, 4), (147,))
```

In [20]:
```
#Decision tree classifier model
model_tree=tree.DecisionTreeClassifier()
model_tree.fit(x,y)   #training model

predit_tree_setosa=model_tree.predict(test_setosa)
print("Decsion tree predicts {} for test setosa".format(iris.target_names[pred
it_tree_setosa]))
```

```
Decsion tree predicts ['setosa'] for test setosa
```

In [22]:
```
predit_tree_vergnica=model_tree.predict(test_vergnica)
print("Decsion tree predicts {} for test verginca".format(iris.target_names[pr
edit_tree_vergnica]))
predit_tree_vertosa=model_tree.predict(test_vertosa)
print("Decsion tree predicts {} for test vertosa".format(iris.target_names[pre
dit_tree_vertosa]))
```

```
Decsion tree predicts ['virginica'] for test verginca
Decsion tree predicts ['versicolor'] for test vertosa
```

Prediction using only two features

In [28]:
```
Y=df['target']
X=df[["petal length (cm)", "petal width (cm)"]]

from sklearn import tree,metrics,model_selection,preprocessing
X_train,X_test,Y_train,Y_test=model_selection.train_test_split(X,Y,test_size=
0.3,random_state=0)
```

In [29]:
```
X_train.shape,X_test.shape,Y_train.shape,Y_test.shape
```

Out[29]:
```
((105, 2), (45, 2), (105,), (45,))
```

In [30]:
```
dtree=tree.DecisionTreeClassifier(criterion='entropy',max_depth=3,random_state
=0)
dtree.fit(X_train,Y_train)
ypred=dtree.predict(X_test)
```

In [31]:
```
dtree
```

Out[31]:
```
DecisionTreeClassifier(class_weight=None, criterion='entropy', max_depth=3,
                       max_features=None, max_leaf_nodes=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, presort=False,
                       random_state=0, splitter='best')
```

In [34]:
```
Y_test,ypred
```

Out[34]:
```
(114    2.0
 62     1.0
 33     0.0
 107    2.0
 7      0.0
 100    2.0
 40     0.0
 86     1.0
 76     1.0
 71     1.0
 134    2.0
 51     1.0
 73     1.0
 54     1.0
 63     1.0
 37     0.0
 78     1.0
 90     1.0
 45     0.0
 16     0.0
 121    2.0
 66     1.0
 24     0.0
 8      0.0
 126    2.0
 22     0.0
 44     0.0
 97     1.0
 93     1.0
 26     0.0
 137    2.0
 84     1.0
 27     0.0
 127    2.0
 132    2.0
 59     1.0
 18     0.0
 83     1.0
 61     1.0
 92     1.0
 112    2.0
 2      0.0
 141    2.0
 43     0.0
 10     0.0
 Name: target, dtype: float64,
 array([2., 1., 0., 2., 0., 2., 0., 1., 1., 1., 2., 1., 1., 1., 1., 0., 1.,
        1., 0., 0., 2., 1., 0., 0., 2., 0., 0., 1., 1., 0., 2., 1., 0., 2.,
        2., 1., 0., 2., 1., 1., 2., 0., 2., 0., 0.]))
```

In [36]:
```
#model performance
miss_classified=(Y_test != ypred).sum()
print("no of samples misclassified ={}".format(miss_classified))
accuracy=metrics.accuracy_score(Y_test,ypred)
print("Accuracy {:.2f}".format(accuracy))
```

```
no of samples misclassified =1
Accuracy 0.98
```

In [39]:
```
conda install -c conda-forge pydotplus
```

```
Collecting package metadata (current_repodata.json): ...working... done
Solving environment: ...working...
The environment is inconsistent, please check the package plan carefully
The following packages are causing the inconsistency:

  - defaults/win-64::anaconda==2019.07=py37_0
  - defaults/win-64::numba==0.44.1=py37hf9181ef_0
  - defaults/win-64::_anaconda_depends==2019.03=py37_0
done

## Package Plan ##

  environment location: C:\Users\vkondala\AppData\Local\Continuum\anaconda3

  added / updated specs:
    - pydotplus


The following packages will be downloaded:

    package                    |            build
    ---------------------------|-----------------
    anaconda-custom            |           py37_1           3 KB
    ca-certificates-2019.6.16  |       hecc5488_0         183 KB  conda-forge
    certifi-2019.6.16          |           py37_1         149 KB  conda-forge
    graphviz-2.38.0            |    h6538335_1011        41.0 MB  conda-forge
    openssl-1.1.1c             |       hfa6e2cd_0         4.7 MB  conda-forge
    pydotplus-2.0.2            |             py_2          23 KB  conda-forge
    tbb-2019.7                 |       he980bc4_0         161 KB  conda-forge
    ------------------------------------------------------------
                                           Total:        46.2 MB

The following NEW packages will be INSTALLED:

  graphviz           conda-forge/win-64::graphviz-2.38.0-h6538335_1011
  pydotplus          conda-forge/noarch::pydotplus-2.0.2-py_2
  tbb                conda-forge/win-64::tbb-2019.7-he980bc4_0

The following packages will be UPDATED:

  ca-certificates    pkgs/main::ca-certificates-2019.5.15-0 --> conda-forge::
ca-certificates-2019.6.16-hecc5488_0
  certifi                                  2019.6.16-py37_0 --> 2019.6.16-py3
7_1

The following packages will be SUPERSEDED by a higher-priority channel:

  openssl              pkgs/main::openssl-1.1.1c-he774522_1 --> conda-forge::
openssl-1.1.1c-hfa6e2cd_0

The following packages will be DOWNGRADED:

  anaconda                                    2019.07-py37_0 --> custom-py37_1



Downloading and Extracting Packages
```

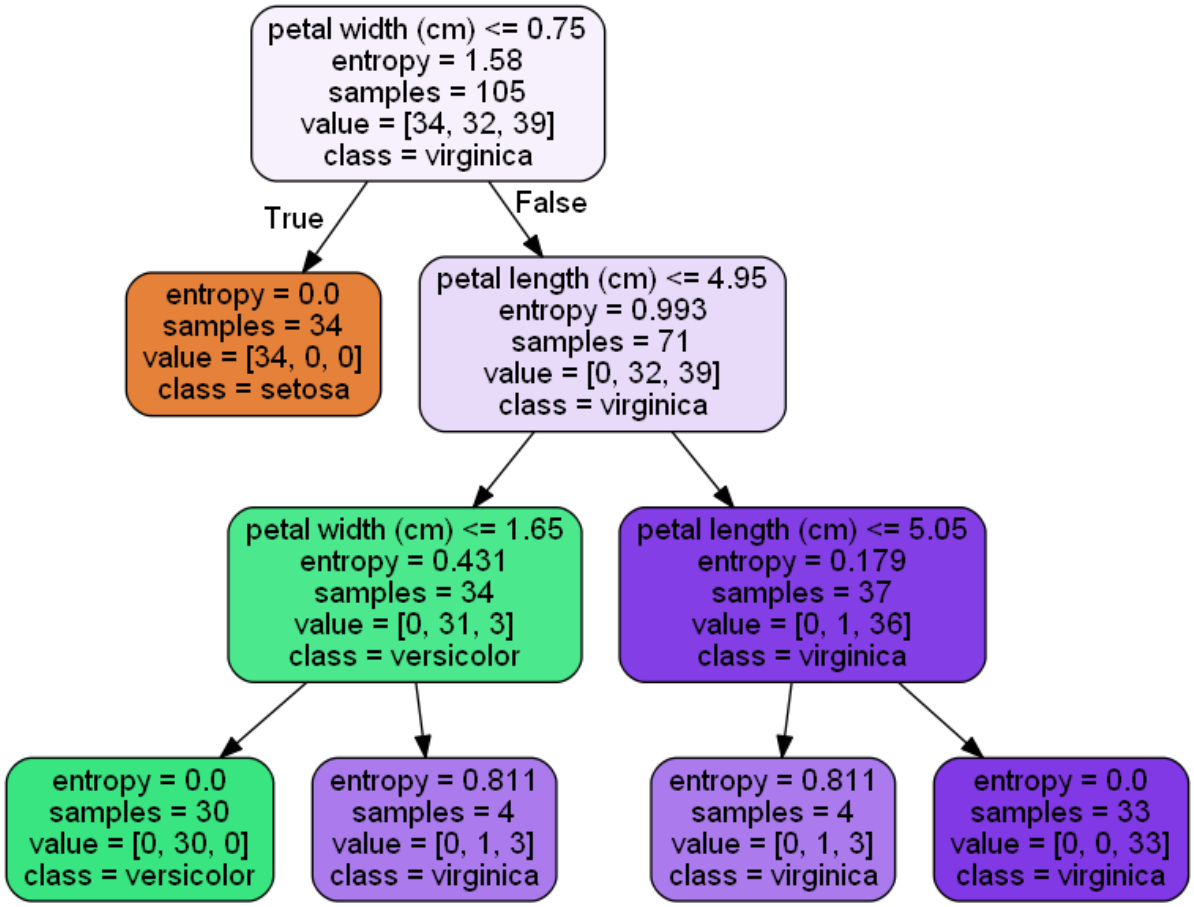
```
graphviz-2.38.0      | 41.0 MB   |            |   0%
graphviz-2.38.0      | 41.0 MB   |            |   0%
graphviz-2.38.0      | 41.0 MB   |            |   0%
graphviz-2.38.0      | 41.0 MB   |            |   1%
graphviz-2.38.0      | 41.0 MB   | 1          |   1%
graphviz-2.38.0      | 41.0 MB   | 1          |   2%
graphviz-2.38.0      | 41.0 MB   | 1          |   2%
graphviz-2.38.0      | 41.0 MB   | 1          |   2%
graphviz-2.38.0      | 41.0 MB   | 2          |   2%
graphviz-2.38.0      | 41.0 MB   | 2          |   2%
graphviz-2.38.0      | 41.0 MB   | 2          |   3%
graphviz-2.38.0      | 41.0 MB   | 2          |   3%
graphviz-2.38.0      | 41.0 MB   | 2          |   3%
graphviz-2.38.0      | 41.0 MB   | 3          |   3%
graphviz-2.38.0      | 41.0 MB   | 3          |   4%
graphviz-2.38.0      | 41.0 MB   | 4          |   4%
graphviz-2.38.0      | 41.0 MB   | 4          |   4%
graphviz-2.38.0      | 41.0 MB   | 4          |   4%
graphviz-2.38.0      | 41.0 MB   | 4          |   5%
graphviz-2.38.0      | 41.0 MB   | 5          |   5%
graphviz-2.38.0      | 41.0 MB   | 5          |   6%
graphviz-2.38.0      | 41.0 MB   | 6          |   6%
graphviz-2.38.0      | 41.0 MB   | 6          |   6%
graphviz-2.38.0      | 41.0 MB   | 6          |   7%
graphviz-2.38.0      | 41.0 MB   | 7          |   7%
graphviz-2.38.0      | 41.0 MB   | 7          |   8%
graphviz-2.38.0      | 41.0 MB   | 8          |   8%
graphviz-2.38.0      | 41.0 MB   | 8          |   9%
graphviz-2.38.0      | 41.0 MB   | 8          |   9%
graphviz-2.38.0      | 41.0 MB   | 9          |   9%
graphviz-2.38.0      | 41.0 MB   | 9          |  10%
graphviz-2.38.0      | 41.0 MB   | #          |  10%
graphviz-2.38.0      | 41.0 MB   | #1         |  11%
graphviz-2.38.0      | 41.0 MB   | #1         |  12%
graphviz-2.38.0      | 41.0 MB   | #2         |  12%
graphviz-2.38.0      | 41.0 MB   | #2         |  13%
graphviz-2.38.0      | 41.0 MB   | #3         |  13%
graphviz-2.38.0      | 41.0 MB   | #3         |  14%
graphviz-2.38.0      | 41.0 MB   | #4         |  15%
graphviz-2.38.0      | 41.0 MB   | #5         |  15%
graphviz-2.38.0      | 41.0 MB   | #6         |  16%
graphviz-2.38.0      | 41.0 MB   | #7         |  17%
graphviz-2.38.0      | 41.0 MB   | #7         |  18%
graphviz-2.38.0      | 41.0 MB   | #8         |  18%
graphviz-2.38.0      | 41.0 MB   | #9         |  19%
graphviz-2.38.0      | 41.0 MB   | #9         |  20%
graphviz-2.38.0      | 41.0 MB   | ##         |  21%
graphviz-2.38.0      | 41.0 MB   | ##1        |  21%
graphviz-2.38.0      | 41.0 MB   | ##2        |  22%
graphviz-2.38.0      | 41.0 MB   | ##2        |  23%
graphviz-2.38.0      | 41.0 MB   | ##3        |  23%
graphviz-2.38.0      | 41.0 MB   | ##3        |  24%
graphviz-2.38.0      | 41.0 MB   | ##4        |  25%
graphviz-2.38.0      | 41.0 MB   | ##5        |  25%
graphviz-2.38.0      | 41.0 MB   | ##5        |  26%
graphviz-2.38.0      | 41.0 MB   | ##6        |  27%
```

```
graphviz-2.38.0      | 41.0 MB   | ##7        |  27%
graphviz-2.38.0      | 41.0 MB   | ##7        |  28%
graphviz-2.38.0      | 41.0 MB   | ##8        |  29%
graphviz-2.38.0      | 41.0 MB   | ##9        |  30%
graphviz-2.38.0      | 41.0 MB   | ###        |  30%
graphviz-2.38.0      | 41.0 MB   | ###        |  31%
graphviz-2.38.0      | 41.0 MB   | ###1       |  31%
graphviz-2.38.0      | 41.0 MB   | ###2       |  32%
graphviz-2.38.0      | 41.0 MB   | ###2       |  33%
graphviz-2.38.0      | 41.0 MB   | ###3       |  33%
graphviz-2.38.0      | 41.0 MB   | ###3       |  34%
graphviz-2.38.0      | 41.0 MB   | ###4       |  34%
graphviz-2.38.0      | 41.0 MB   | ###5       |  35%
graphviz-2.38.0      | 41.0 MB   | ###5       |  36%
graphviz-2.38.0      | 41.0 MB   | ###6       |  37%
graphviz-2.38.0      | 41.0 MB   | ###7       |  37%
graphviz-2.38.0      | 41.0 MB   | ###8       |  38%
graphviz-2.38.0      | 41.0 MB   | ###8       |  39%
graphviz-2.38.0      | 41.0 MB   | ###9       |  39%
graphviz-2.38.0      | 41.0 MB   | ####       |  40%
graphviz-2.38.0      | 41.0 MB   | ####       |  41%
graphviz-2.38.0      | 41.0 MB   | ####1      |  41%
graphviz-2.38.0      | 41.0 MB   | ####2      |  42%
graphviz-2.38.0      | 41.0 MB   | ####2      |  43%
graphviz-2.38.0      | 41.0 MB   | ####3      |  44%
graphviz-2.38.0      | 41.0 MB   | ####4      |  44%
graphviz-2.38.0      | 41.0 MB   | ####5      |  45%
graphviz-2.38.0      | 41.0 MB   | ####5      |  46%
graphviz-2.38.0      | 41.0 MB   | ####6      |  46%
graphviz-2.38.0      | 41.0 MB   | ####6      |  47%
graphviz-2.38.0      | 41.0 MB   | ####7      |  48%
graphviz-2.38.0      | 41.0 MB   | ####8      |  48%
graphviz-2.38.0      | 41.0 MB   | ####8      |  48%
graphviz-2.38.0      | 41.0 MB   | ####9      |  49%
graphviz-2.38.0      | 41.0 MB   | ####9      |  50%
graphviz-2.38.0      | 41.0 MB   | ####9      |  50%
graphviz-2.38.0      | 41.0 MB   | #####      |  50%
graphviz-2.38.0      | 41.0 MB   | #####      |  51%
graphviz-2.38.0      | 41.0 MB   | #####1     |  51%
graphviz-2.38.0      | 41.0 MB   | #####1     |  51%
graphviz-2.38.0      | 41.0 MB   | #####1     |  52%
graphviz-2.38.0      | 41.0 MB   | #####1     |  52%
graphviz-2.38.0      | 41.0 MB   | #####2     |  52%
graphviz-2.38.0      | 41.0 MB   | #####2     |  53%
graphviz-2.38.0      | 41.0 MB   | #####3     |  53%
graphviz-2.38.0      | 41.0 MB   | #####3     |  54%
graphviz-2.38.0      | 41.0 MB   | #####4     |  55%
graphviz-2.38.0      | 41.0 MB   | #####5     |  55%
graphviz-2.38.0      | 41.0 MB   | #####5     |  56%
graphviz-2.38.0      | 41.0 MB   | #####6     |  56%
graphviz-2.38.0      | 41.0 MB   | #####6     |  57%
graphviz-2.38.0      | 41.0 MB   | #####7     |  57%
graphviz-2.38.0      | 41.0 MB   | #####7     |  58%
graphviz-2.38.0      | 41.0 MB   | #####8     |  59%
graphviz-2.38.0      | 41.0 MB   | #####9     |  59%
graphviz-2.38.0      | 41.0 MB   | ######     |  60%
graphviz-2.38.0      | 41.0 MB   | ######1    |  61%
```

```
graphviz-2.38.0      | 41.0 MB   | ######2    |  62%
graphviz-2.38.0      | 41.0 MB   | ######2    |  63%
graphviz-2.38.0      | 41.0 MB   | ######3    |  64%
graphviz-2.38.0      | 41.0 MB   | ######4    |  65%
graphviz-2.38.0      | 41.0 MB   | ######5    |  66%
graphviz-2.38.0      | 41.0 MB   | ######6    |  66%
graphviz-2.38.0      | 41.0 MB   | ######6    |  67%
graphviz-2.38.0      | 41.0 MB   | ######7    |  67%
graphviz-2.38.0      | 41.0 MB   | ######8    |  68%
graphviz-2.38.0      | 41.0 MB   | ######8    |  69%
graphviz-2.38.0      | 41.0 MB   | ######9    |  69%
graphviz-2.38.0      | 41.0 MB   | ######9    |  70%
graphviz-2.38.0      | 41.0 MB   | #######    |  70%
graphviz-2.38.0      | 41.0 MB   | #######1   |  71%
graphviz-2.38.0      | 41.0 MB   | #######2   |  72%
graphviz-2.38.0      | 41.0 MB   | #######2   |  73%
graphviz-2.38.0      | 41.0 MB   | #######3   |  74%
graphviz-2.38.0      | 41.0 MB   | #######4   |  74%
graphviz-2.38.0      | 41.0 MB   | #######4   |  75%
graphviz-2.38.0      | 41.0 MB   | #######5   |  75%
graphviz-2.38.0      | 41.0 MB   | #######6   |  77%
graphviz-2.38.0      | 41.0 MB   | #######7   |  77%
graphviz-2.38.0      | 41.0 MB   | #######8   |  78%
graphviz-2.38.0      | 41.0 MB   | #######8   |  79%
graphviz-2.38.0      | 41.0 MB   | #######9   |  80%
graphviz-2.38.0      | 41.0 MB   | ########   |  80%
graphviz-2.38.0      | 41.0 MB   | ########1  |  81%
graphviz-2.38.0      | 41.0 MB   | ########2  |  82%
graphviz-2.38.0      | 41.0 MB   | ########3  |  83%
graphviz-2.38.0      | 41.0 MB   | ########4  |  84%
graphviz-2.38.0      | 41.0 MB   | ########5  |  85%
graphviz-2.38.0      | 41.0 MB   | ########6  |  86%
graphviz-2.38.0      | 41.0 MB   | ########6  |  87%
graphviz-2.38.0      | 41.0 MB   | ########7  |  87%
graphviz-2.38.0      | 41.0 MB   | ########8  |  88%
graphviz-2.38.0      | 41.0 MB   | ########8  |  89%
graphviz-2.38.0      | 41.0 MB   | ########9  |  90%
graphviz-2.38.0      | 41.0 MB   | #########  |  90%
graphviz-2.38.0      | 41.0 MB   | #########1 |  91%
graphviz-2.38.0      | 41.0 MB   | #########1 |  92%
graphviz-2.38.0      | 41.0 MB   | #########2 |  92%
graphviz-2.38.0      | 41.0 MB   | #########2 |  93%
graphviz-2.38.0      | 41.0 MB   | #########3 |  93%
graphviz-2.38.0      | 41.0 MB   | #########3 |  94%
graphviz-2.38.0      | 41.0 MB   | #########5 |  95%
graphviz-2.38.0      | 41.0 MB   | #########5 |  96%
graphviz-2.38.0      | 41.0 MB   | #########6 |  96%
graphviz-2.38.0      | 41.0 MB   | #########7 |  97%
graphviz-2.38.0      | 41.0 MB   | #########8 |  98%
graphviz-2.38.0      | 41.0 MB   | #########8 |  99%
graphviz-2.38.0      | 41.0 MB   | #########9 | 100%
graphviz-2.38.0      | 41.0 MB   | ########## | 100%

tbb-2019.7           | 161 KB    |            |   0%
tbb-2019.7           | 161 KB    | ######9    |  70%
tbb-2019.7           | 161 KB    | ########## | 100%
```

```
pydotplus-2.0.2      | 23 KB     |            |   0%
pydotplus-2.0.2      | 23 KB     | ########## | 100%

ca-certificates-2019 | 183 KB    |            |   0%
ca-certificates-2019 | 183 KB    | ########## | 100%

openssl-1.1.1c       | 4.7 MB    |            |   0%
openssl-1.1.1c       | 4.7 MB    | 3          |   4%
openssl-1.1.1c       | 4.7 MB    | 9          |   9%
openssl-1.1.1c       | 4.7 MB    | #7         |  18%
openssl-1.1.1c       | 4.7 MB    | ##6        |  27%
openssl-1.1.1c       | 4.7 MB    | ###8       |  39%
openssl-1.1.1c       | 4.7 MB    | ####8      |  49%
openssl-1.1.1c       | 4.7 MB    | #####6     |  56%
openssl-1.1.1c       | 4.7 MB    | ######3    |  64%
openssl-1.1.1c       | 4.7 MB    | #######6   |  76%
openssl-1.1.1c       | 4.7 MB    | ########6  |  86%
openssl-1.1.1c       | 4.7 MB    | #########4 |  95%
openssl-1.1.1c       | 4.7 MB    | ########## | 100%

anaconda-custom      | 3 KB      |            |   0%
anaconda-custom      | 3 KB      | ########## | 100%

certifi-2019.6.16    | 149 KB    |            |   0%
certifi-2019.6.16    | 149 KB    | #######5   |  75%
certifi-2019.6.16    | 149 KB    | ########## | 100%
Preparing transaction: ...working... done
Verifying transaction: ...working... done
Executing transaction: ...working... done

Note: you may need to restart the kernel to use updated packages.
```

In [40]:
```
from IPython.display import Image,display
import numpy as np, pandas as pd, matplotlib.pyplot as plt, pydotplus
dot_data=tree.export_graphviz(dtree,out_file=None,filled=True,rounded=True,fea
ture_names=["petal length (cm)", "petal width (cm)"],
                              class_names=['setosa','versicolor','virginica'])

graph=pydotplus.graph_from_dot_data(dot_data)
display(Image(graph.create_png()))
```



In [ ]: