Confusion_Matrix

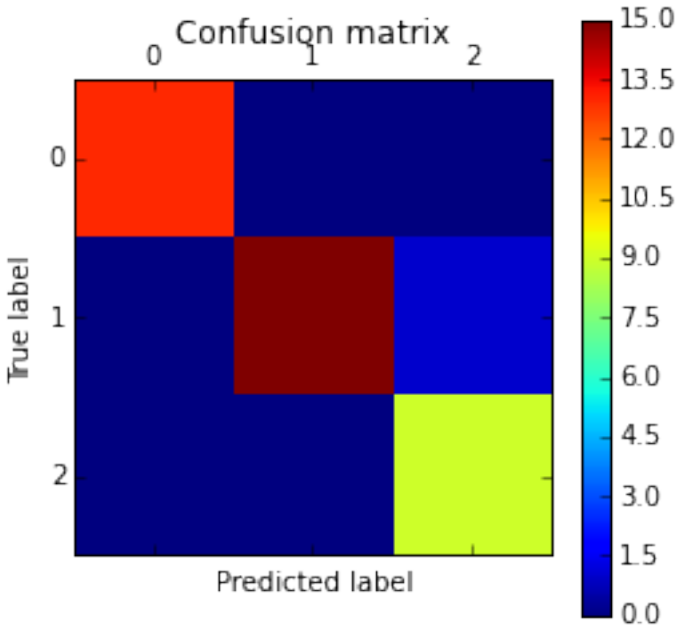
March 27, 2015

In [1]:
```
#Author: Muhammed Khan
```

In [3]:
```
%matplotlib inline
#http://scikit-learn.org/stable/auto_examples/plot_confusion_matrix.html
#Used to determine the quality of the output of a classifier.
print(__doc__)

from sklearn import svm, datasets
from sklearn.cross_validation import train_test_split
from sklearn.metrics import confusion_matrix

import matplotlib.pyplot as plt

# import some data to play with
iris = datasets.load_iris()
X = iris.data
y = iris.target #Attribute we wmat to target

# Split the data into a training set and a test set
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=0)

# Run classifier
classifier = svm.SVC(kernel='linear') #Running a linear support vector machine
y_pred = classifier.fit(X_train, y_train).predict(X_test)

# Compute confusion matrix
cm = confusion_matrix(y_test, y_pred)

print(cm)

# Show confusion matrix in a separate window
plt.matshow(cm)
plt.title('Confusion matrix')
plt.colorbar()
plt.ylabel('True label')
plt.xlabel('Predicted label')
plt.show()
#Veriscolor(15) followed by setosa (13)
#plants were most accurately classified in the irish dataset Which on the whole was
#very well classified  The largest amount of misclassifications occured when the plant was iden
#virginica but was really versicolor
```

```
Automatically created module for IPython interactive environment
[[13  0  0]
```

```
[ 0 15  1]
[ 0  0  9]]
```


Confusion matrix
0
1
2
0
1
2
True label
Predicted label
15.0
13.5
12.0
10.5
9.0
7.5
6.0
4.5
3.0
1.5
0.0