In [17]:
```
import numpy as np
import pandas as pd
from sklearn import cross_validation
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
spambase=pd.read_csv('C:\Users\SGGS3\Documents\spambase.data')
list1=['f1','f2','f3','f4','f5','f6','f7','f8','f9','f10','f11','f12','f13','f14'
list2=['class']
X=spambase[list1]
y=spambase[list2]
X=np.array(X)
y=np.array(y)
X_train,X_test,y_train,y_test=cross_validation.train_test_split(X,y,test_size=0.3
gnb.fit(X_train,y_train)
acc=gnb.score(X_test,y_test)
print "Accuracy of Naive-Bayes is :- ",acc
```

```
C:\Anaconda2\lib\site-packages\sklearn\utils\validation.py:515: DataConversionW
arning: A column-vector y was passed when a 1d array was expected. Please chang
e the shape of y to (n_samples, ), for example using ravel().
  y = column_or_1d(y, warn=True)

Accuracy of Naive-Bayes is :-  0.821868211441
```

In [16]:
```
print(len(spambase))
print(len(X_train))
print(len(y_train))
print(len(X_test))
print(len(y_test))
```

```
4601
3220
3220
1381
1381
```