In [2]:
```
#Step1:)Import required libraries
import pandas
from pandas.plotting import scatter_matrix
import matplotlib.pyplot as plt
from sklearn import model_selection
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
from sklearn.linear_model import LogisticRegression  #LR
from sklearn.tree import DecisionTreeClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
from sklearn.naive_bayes import GaussianNB #NB
from sklearn.svm import SVC
```

In [5]:
```
#step2)Loading dataset
url = "https://raw.githubusercontent.com/jbrownlee/Datasets/master/iris.csv"
names = ['sepal-length', 'sepal-width', 'petal-length', 'petal-width', 'clas
s']
dataset = pandas.read_csv(url, names=names) #(url,name=var where labels are)
dataset.head()
```

Out[5]:

	sepal-length	sepal-width	petal-length	petal-width	class
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

In [4]:
```
#shape->no of rows and columns
print(dataset.shape)
```

```
(150, 5)
```

In [5]:
```
print(dataset.head(50))
```

```
    sepal-length  sepal-width  petal-length  petal-width        class
0            5.1          3.5           1.4          0.2  Iris-setosa
1            4.9          3.0           1.4          0.2  Iris-setosa
2            4.7          3.2           1.3          0.2  Iris-setosa
3            4.6          3.1           1.5          0.2  Iris-setosa
4            5.0          3.6           1.4          0.2  Iris-setosa
5            5.4          3.9           1.7          0.4  Iris-setosa
6            4.6          3.4           1.4          0.3  Iris-setosa
7            5.0          3.4           1.5          0.2  Iris-setosa
8            4.4          2.9           1.4          0.2  Iris-setosa
9            4.9          3.1           1.5          0.1  Iris-setosa
10           5.4          3.7           1.5          0.2  Iris-setosa
11           4.8          3.4           1.6          0.2  Iris-setosa
12           4.8          3.0           1.4          0.1  Iris-setosa
13           4.3          3.0           1.1          0.1  Iris-setosa
14           5.8          4.0           1.2          0.2  Iris-setosa
15           5.7          4.4           1.5          0.4  Iris-setosa
16           5.4          3.9           1.3          0.4  Iris-setosa
17           5.1          3.5           1.4          0.3  Iris-setosa
18           5.7          3.8           1.7          0.3  Iris-setosa
19           5.1          3.8           1.5          0.3  Iris-setosa
20           5.4          3.4           1.7          0.2  Iris-setosa
21           5.1          3.7           1.5          0.4  Iris-setosa
22           4.6          3.6           1.0          0.2  Iris-setosa
23           5.1          3.3           1.7          0.5  Iris-setosa
24           4.8          3.4           1.9          0.2  Iris-setosa
25           5.0          3.0           1.6          0.2  Iris-setosa
26           5.0          3.4           1.6          0.4  Iris-setosa
27           5.2          3.5           1.5          0.2  Iris-setosa
28           5.2          3.4           1.4          0.2  Iris-setosa
29           4.7          3.2           1.6          0.2  Iris-setosa
30           4.8          3.1           1.6          0.2  Iris-setosa
31           5.4          3.4           1.5          0.4  Iris-setosa
32           5.2          4.1           1.5          0.1  Iris-setosa
33           5.5          4.2           1.4          0.2  Iris-setosa
34           4.9          3.1           1.5          0.1  Iris-setosa
35           5.0          3.2           1.2          0.2  Iris-setosa
36           5.5          3.5           1.3          0.2  Iris-setosa
37           4.9          3.1           1.5          0.1  Iris-setosa
38           4.4          3.0           1.3          0.2  Iris-setosa
39           5.1          3.4           1.5          0.2  Iris-setosa
40           5.0          3.5           1.3          0.3  Iris-setosa
41           4.5          2.3           1.3          0.3  Iris-setosa
42           4.4          3.2           1.3          0.2  Iris-setosa
43           5.0          3.5           1.6          0.6  Iris-setosa
44           5.1          3.8           1.9          0.4  Iris-setosa
45           4.8          3.0           1.4          0.3  Iris-setosa
46           5.1          3.8           1.6          0.2  Iris-setosa
47           4.6          3.2           1.4          0.2  Iris-setosa
48           5.3          3.7           1.5          0.2  Iris-setosa
49           5.0          3.3           1.4          0.2  Iris-setosa
```

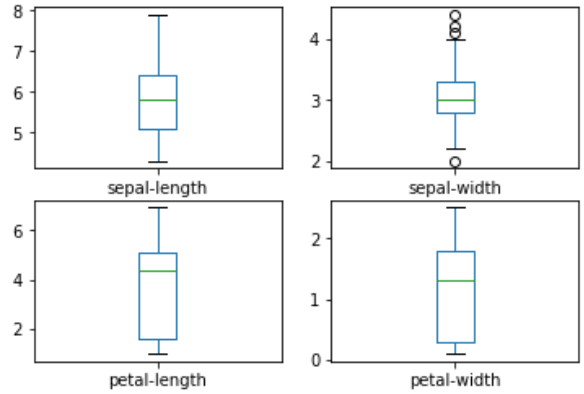
In [8]:
```
#Step3:Summerization of dataset
#descriptions
print(dataset.describe())
```

```
       sepal-length  sepal-width  petal-length  petal-width
count    150.000000   150.000000    150.000000   150.000000
mean       5.843333     3.054000      3.758667     1.198667
std        0.828066     0.433594      1.764420     0.763161
min        4.300000     2.000000      1.000000     0.100000
25%        5.100000     2.800000      1.600000     0.300000
50%        5.800000     3.000000      4.350000     1.300000
75%        6.400000     3.300000      5.100000     1.800000
max        7.900000     4.400000      6.900000     2.500000
```

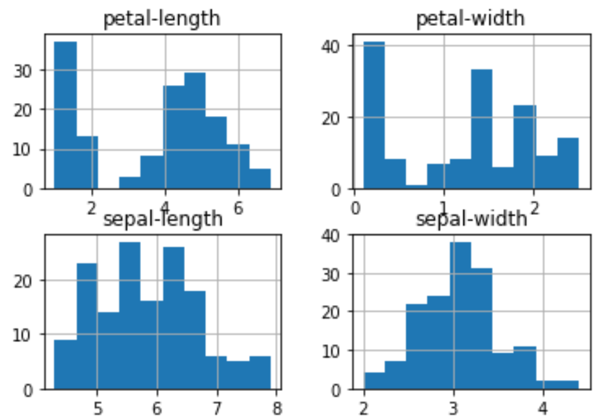
In [7]:
```
#class distribution
print(dataset.groupby('class').size())
```

```
class
Iris-setosa        50
Iris-versicolor    50
Iris-virginica     50
dtype: int64
```

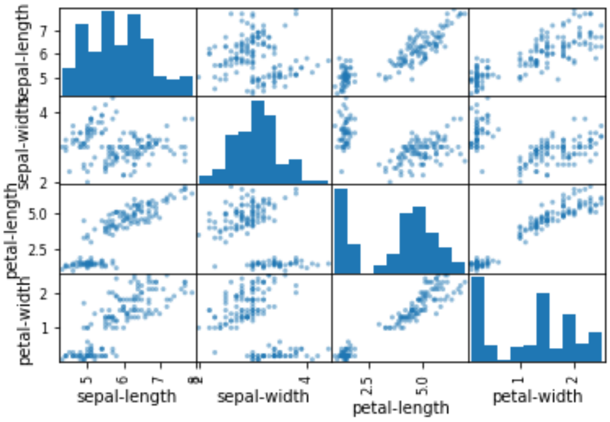
In [8]:
```
# box and whisker plots
#univariate_plot  (diff attr->diff ploting)
dataset.plot(kind='box', subplots=True, layout=(2,2), sharex=T, sharey=False
)
plt.show()
```



In [9]:
```
#Step:4) Visualization of dataset
#histograms #univariate_plot
dataset.hist()
plt.show()
```



In [11]:
```
#multivariate_plot
scatter_matrix(dataset)
plt.show()
```



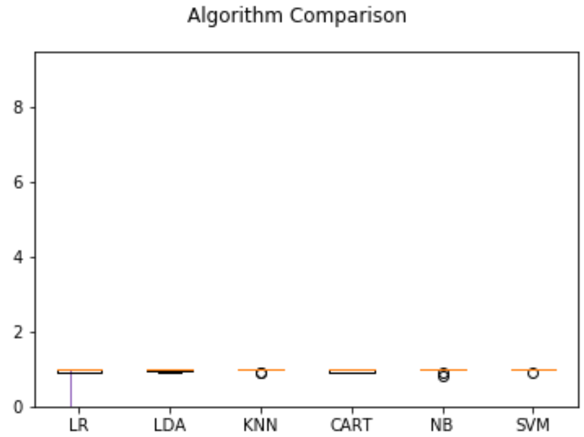
In [9]:
```
#diving dataset for training(with-label) and testing(without-label) datasets
# Split-out validation dataset
array = dataset.values
X = array[:,0:4] #training-data column 1 to 4
Y = array[:,4] #training-answer column 5
validation_size = 0.20   #20% of dataset
seed = 7  #value for random number generation
X_train, X_validation, Y_train, Y_validation = model_selection.train_test_sp
lit
(X, Y, test_size=validation_size, random_state=seed #based on seed)
```

In [10]:
```
#Step4)Evaluating algorithms
seed = 7
scoring = 'accuracy'
models = []
models.append(('LR', LogisticRegression()))
models.append(('LDA', LinearDiscriminantAnalysis()))
models.append(('KNN', KNeighborsClassifier()))
models.append(('CART', DecisionTreeClassifier()))
models.append(('NB', GaussianNB()))
models.append(('SVM', SVC()))
# evaluate each model in turn
results = []
names = []
for name, model in models:
        kfold = model_selection.KFold(n_splits=10, random_state=seed) #divid
ing 120 in 10 10 epochs
        cv_results = model_selection.cross_val_score(model, X_train, Y_train
, cv=kfold, scoring=scoring)
        results.append(cv_results)
        names.append(name)
        msg = "%s: %f (%f)" % (name, cv_results.mean(), cv_results.std())
        print(msg)
```

```
LR: 0.966667 (0.040825)
LDA: 0.975000 (0.038188)
KNN: 0.983333 (0.033333)
CART: 0.975000 (0.038188)
NB: 0.975000 (0.053359)
SVM: 0.991667 (0.025000)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
33: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify
a solver to silence this warning.
  FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
60: FutureWarning: Default multi_class will be changed to 'auto' in 0.22. Spe
cify the multi_class option to silence this warning.
  "this warning.", FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
33: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify
a solver to silence this warning.
  FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
60: FutureWarning: Default multi_class will be changed to 'auto' in 0.22. Spe
cify the multi_class option to silence this warning.
  "this warning.", FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
33: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify
a solver to silence this warning.
  FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
60: FutureWarning: Default multi_class will be changed to 'auto' in 0.22. Spe
cify the multi_class option to silence this warning.
  "this warning.", FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
33: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify
a solver to silence this warning.
  FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
60: FutureWarning: Default multi_class will be changed to 'auto' in 0.22. Spe
cify the multi_class option to silence this warning.
  "this warning.", FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
33: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify
a solver to silence this warning.
  FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
60: FutureWarning: Default multi_class will be changed to 'auto' in 0.22. Spe
cify the multi_class option to silence this warning.
  "this warning.", FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
33: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify
a solver to silence this warning.
  FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
60: FutureWarning: Default multi_class will be changed to 'auto' in 0.22. Spe
cify the multi_class option to silence this warning.
  "this warning.", FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
33: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify
a solver to silence this warning.
  FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
60: FutureWarning: Default multi_class will be changed to 'auto' in 0.22. Spe
cify the multi_class option to silence this warning.
  "this warning.", FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
33: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify
a solver to silence this warning.
  FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
60: FutureWarning: Default multi_class will be changed to 'auto' in 0.22. Spe
cify the multi_class option to silence this warning.
  "this warning.", FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:4
33: FutureWarning: Default solver will be changed to 'lbfgs' in 0.22. Specify
a solver to silence this warning.
  FutureWarning)
```

In [12]:
```
# Compare Algorithms
fig = plt.figure()
fig.suptitle('Algorithm Comparison')
#plt.hist(results)
ax = fig.add_subplot(111)
plt.boxplot(results)
plt.hist(results)
ax.set_xticklabels(names)
plt.show()
```



In [14]:
```
#Step:5) Making predictions
# Make predictions on validation dataset
knn = KNeighborsClassifier()
knn.fit(X_train, Y_train) #fitting values for generalized model
predictions = knn.predict(X_validation) #prediction on remaining 20% of data
set that is in X_Validation
print("accurracy score",accuracy_score(Y_validation, predictions)) #comparin
g with y-valid for accuracy
print("confusion matrix \n",confusion_matrix(Y_validation, predictions))
#print(classification_report(Y_validation, predictions))
```

```
accurracy score 0.9
confusion matrix
 [[ 7  0  0]
 [ 0 11  1]
 [ 0  2  9]]
```

In [ ]:
```
#MATRIX:
#        classes
#  predi ----------
#  cted  |
#  data  |
#7 are from class 1 (iris-sentosa) 7-right predicted in class 1
#12 are from class 2 (iris-versicolor) 11 in class 2 and 1 false predicted i
n class 3
#11 are from class 3 (iris-) 2 false predicted in class 2 and 9 is true pred
iction of class 3
```

In [ ]:
```

```