In [84]:

```
import numpy as np
import pandas as pd
from sklearn import datasets
from sklearn.datasets import load_iris
from sklearn import model_selection
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.model_selection import cross_val_score,KFold,RepeatedKFold
from sklearn.ensemble import AdaBoostClassifier
from sklearn.ensemble import GradientBoostingClassifier
```

Importing the Dataset

In [77]:

```
df = pd.DataFrame(load_iris()['data'],
columns=load_iris()['feature_names'])
df['target'] = load_iris()['target']
df.head()
```

Out[77]:

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

Splitting the dataset

In [78]:

```
X = df.drop(['target'],axis =1)
y = df['target']
X_train,X_test,y_train,y_test = train_test_split(X,y,test_size=0.3, random_state = 42)
```

Fitting Decision tree

In [79]:

```
from sklearn.tree import DecisionTreeClassifier
tree = DecisionTreeClassifier(criterion = "entropy",max_depth = 3)
tree.fit(X_train,y_train)
print(tree.score(X_train,y_train))
y_pred=tree.predict(X_test)

print(f"Decision Tree Classifier testing accuracy: {accuracy_score(y_test,y_pred)}")
```

```
0.9523809523809523
Decision Tree Classifier testing accuracy: 0.9777777777777777
```

```
Fitting the adaboostclassifier
```

In [80]:

```
adb = AdaBoostClassifier(tree,n_estimators = 10 , learning_rate = 0.1)
adb.fit(X_train,y_train)
print("score",adb.score(X_train,y_train))
y_pred = adb.predict(X_test)
print(f"AdaBoost Classifier testing accuracy: {accuracy_score(y_test,y_pred)}")
```

```
score 1.0
AdaBoost Classifier testing accuracy: 1.0
```

In [89]:

```
ada_cv = cross_val_score(adb,X_train,y_train,cv = 4)
print(f"AdaBoost Classifier Cross Validation Score:{ada_cv.mean()}")
```

```
AdaBoost Classifier Cross Validation Score:0.8764245014245015
```

In [90]:

```
gda_cv = cross_val_score(gdb,X_train,y_train,cv = KFold(n_splits = 4) )
print(f"AdaBoost Classifier KFold Cross Validation Score:{gda_cv.mean()}")
```

```
AdaBoost Classifier KFold Cross Validation Score:0.9234330484330484
```

In [88]:

```
gda_cv = cross_val_score(gdb,X_train,y_train,cv = RepeatedKFold(n_splits = 4) )
print(f"AdaBoost Classifier RepeatedKFold Cross Validation Score:{gda_cv.mean()}")
```

```
Gradient Boosting Classifier RepeatedKFold Cross Validation Score:0.91438746
43874643
```

```
Fitting the Gradient Boosting Classifier
```

In [82]:

```
gdb = GradientBoostingClassifier(init = tree,n_estimators = 10 , learning_rate = 0.1)
gdb.fit(X_train,y_train)
print("score",gdb.score(X_train,y_train))
y_pred = gdb.predict(X_test)
print(f"GradientBoost Classifier testing accuracy: {accuracy_score(y_test,y_pred)}")
```

```
score 1.0
GradientBoost Classifier testing accuracy: 0.9555555555555556
```

In [83]:

```
gda_cv = cross_val_score(gdb,X_train,y_train,cv = 4)
print(f"Gradient Boosting Classifier Cross Validation Score:{gda_cv.mean()}")
```

```
Gradient Boosting Classifier Cross Validation Score:0.9045584045584045
```

In [85]:

```
gda_cv = cross_val_score(gdb,X_train,y_train,cv = KFold(n_splits = 4) )
print(f"Gradient Boosting Classifier KFold Cross Validation Score:{gda_cv.mean()}")
```

```
Gradient Boosting Classifier KFold Cross Validation Score:0.9234330484330484
```

In [86]:

```
gda_cv = cross_val_score(gdb,X_train,y_train,cv = RepeatedKFold(n_splits = 4) )
print(f"Gradient Boosting Classifier RepeatedKFold Cross Validation Score:{gda_cv.mean()}")
```

```
Gradient Boosting Classifier KFold Cross Validation Score:0.9152777777777779
```

1)Since both the adaboost and gradient boosting models gives score as 1.0 for training it is obvious that the models are overfitting.Seqential learning models tends to overfit.

2)The testing accuracy for adaboost is 1.0 and for gradient boosting the accuracy is 0.9555555.the metrics used is accuracy_score.

3)We have done the cross-validation using cv,kfold,RepeatedKFold.The scores are

	Adaboost	GradientBoosting
CV	0.8764245014245015	0.9045584045584045
KFold	0.9234330484330484	0.9234330484330484
RepeatedKFold	0.9143874643874643	0.9152777777777779

These cross validation scores tells us that the accuracies for Trainig and Testing are not accurate.