Question 1 : Implement the Bagging based Ensemble Model using CART (Classification and Regression Trees) as base learners. No. of base learners = 100. Use cross validation as the model estimation method.

```
import pandas
from sklearn import model_selection
from sklearn.ensemble import BaggingClassifier
from sklearn.tree import DecisionTreeClassifier
url = "https://raw.githubusercontent.com/jbrownlee/Datasets/master/pima-indians-diabetes.data.csv"
names = ['preg','plas','pres','skin','test','mass','pedi','age','class']
df=pandas.read_csv(url,names=names)
df.shape
```

```
(768, 9)
```

```
array=df.values
X = array[:,0:8]
Y = array[:,8]
Kfold = model_selection.KFold(n_splits=10, random_state=None)
cart = DecisionTreeClassifier()
num_trees = 100
model = BaggingClassifier(base_estimator=cart, n_estimators=num_trees,random_state=7)
results=model_selection.cross_val_score(model,X,Y,cv = Kfold)
average_accuracy = sum(results)/len(results)
print("Average Accuracy is ", average_accuracy)
```

```
Average Accuracy is  0.7720437457279563
```

Question 2 : Implement the AdaBoost Ensemble model for classification using 10 base learners and cross validation.

```
import pandas
from sklearn import model_selection
from sklearn.ensemble import AdaBoostClassifier
from sklearn.tree import DecisionTreeClassifier
url = "https://raw.githubusercontent.com/jbrownlee/Datasets/master/pima-indians-diabetes.data.csv"
names = ['preg','plas','pres','skin','test','mass','pedi','age','class']
df=pandas.read_csv(url,names=names)
df.shape
```

```
(768, 9)
```

```
array=df.values
X = array[:,0:8]
Y = array[:,8]
Kfold = model_selection.KFold(n_splits=10, random_state=None)
cart = DecisionTreeClassifier()
num_trees = 100
model = AdaBoostClassifier(base_estimator=cart, n_estimators=num_trees,random_state=7)
results=model_selection.cross_val_score(model,X,Y,cv = Kfold)
average_accuracy = sum(results)/len(results)
print("Average Accuracy is ", average_accuracy)
```

```
Average Accuracy is  0.6964969241285031
```

Question 3: (Same as Question 1) Implement the Bagging based Ensemble Model using k-Nearest Neighbor Classifier as base learners. No. of base learners = 100. Use cross validation as the model estimation method.

```
import pandas
from sklearn import model_selection
from sklearn.ensemble import BaggingClassifier
from sklearn.neighbors import KNeighborsClassifier
url="https://raw.githubusercontent.com/jbrownlee/Datasets/master/pima-indians-diabetes.data.csv"
names=['preg','plas','pres','skin','test','mass','pedi','age','class']
df=pandas.read_csv(url,names=names)
df.shape
```

```
(768, 9)
```

```
array=df.values
X = array[:,0:8]
Y = array[:,8]
Kfold = model_selection.KFold(n_splits=10, random_state=None)
cart = KNeighborsClassifier()
num_trees = 100
model =BaggingClassifier(base_estimator=cart, n_estimators=num_trees,random_state=7)
results=model_selection.cross_val_score(model,X,Y,cv = Kfold)
average_accuracy = sum(results)/len(results)
print("Average Accuracy is ", average_accuracy)
```

```
Average Accuracy is  0.7356459330143541
```