In [1]:
```
### AUTHOR : PUSHPAK VIJAY KATKHEDE
### AI 539 - ML Challenges
### Assignment 2: Give Your Models a Grade
### DATE : 02/04/2023

#importing required modules
import pandas as pd
import numpy as np
from sklearn import tree
from sklearn.ensemble import RandomForestClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.neural_network import MLPClassifier
from sklearn.model_selection import StratifiedKFold, GroupKFold, KFold, StratifiedGrou
from tabulate import tabulate
from sklearn.dummy import DummyClassifier

#reading in the required files via pandas
data = pd.read_csv('activity-dev.csv')
test = pd.read_csv("activity-heldout.csv")

#defining the feature labels and target label
X = data[['G_front', 'G_vert', 'G_lat', 'ant_id', 'RSSI', 'phase', 'freq']]
y = data['activity']

## A : train test split

#defining classifiers
clf_tree = tree.DecisionTreeClassifier(random_state=9)
clf_ranFor = RandomForestClassifier(max_depth=2, random_state=9)
clf_knnc =KNeighborsClassifier(n_neighbors=3)
clf_mlp = MLPClassifier(random_state=9, max_iter=300)

#seperating data into seperate train and test dataframes
X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.8, random_state

#Training all the classifiers and calculating accuracy on each of the classifier
clf_tree.fit(X_train, y_train)
accuracy_4_a_tree = clf_tree.score(X_test, y_test)

clf_ranFor.fit(X_train, y_train)
accuracy_4_a_ranFor = clf_ranFor.score(X_test, y_test)

clf_knnc.fit(X_train, y_train)
accuracy_4_a_knnc = clf_knnc.score(X_test, y_test)

clf_mlp.fit(X_train, y_train)
accuracy_4_a_mlp = clf_mlp.score(X_test, y_test)



## B : 10 fold validation

#defining classifiers
clf_tree_kfold = tree.DecisionTreeClassifier(random_state=9)
clf_ranFor_kfold = RandomForestClassifier(max_depth=2, random_state=9)
clf_knnc_kfold =  KNeighborsClassifier(n_neighbors=3)
clf_mlp_kfold = MLPClassifier(random_state=9, max_iter=300)
```

```
#defining kfold object to define the evaulation strategy
kf = KFold(n_splits=10,  shuffle=True, random_state=9)

#Assigning the evaluation strategy to the classifier and then averaging the mean of th
kfold_scores_tree = cross_val_score(clf_tree_kfold, X, y, scoring='accuracy', cv=kf, r
kfold_score_tree = np.mean(kfold_scores_tree)

kfold_scores_ranFor = cross_val_score(clf_ranFor_kfold, X, y, scoring='accuracy', cv=k
kfold_score_ranFor = np.mean(kfold_scores_ranFor)

kfold_scores_knnc = cross_val_score(clf_knnc_kfold, X, y, scoring='accuracy', cv=kf, r
kfold_score_knnc = np.mean(kfold_scores_knnc)

kfold_scores_mlp = cross_val_score(clf_mlp_kfold, X, y, scoring='accuracy', cv=kf, n_j
kfold_score_mlp = np.mean(kfold_scores_mlp)


## C Stratified 10 fold

#defining classifiers
clf_tree_skfold = tree.DecisionTreeClassifier(random_state=9)
clf_ranFor_skfold = RandomForestClassifier(max_depth=2, random_state=9)
clf_knnc_skfold =  KNeighborsClassifier(n_neighbors=3)
clf_mlp_skfold = MLPClassifier(random_state=9, max_iter=300)

#defining kfold object to define the evaulation strategy
skf = StratifiedKFold(n_splits=10, shuffle=True, random_state=9)

#Assigning the evaluation strategy to the classifier and then averging the mean of the
skfold_scores_tree = cross_val_score(clf_tree_skfold, X, y, scoring='accuracy', cv=skf
skfold_score_tree = np.mean(skfold_scores_tree)

skfold_scores_ranFor = cross_val_score(clf_ranFor_skfold, X, y, scoring='accuracy', cv
skfold_score_ranFor = np.mean(skfold_scores_ranFor)

skfold_scores_knnc = cross_val_score(clf_knnc_skfold, X, y, scoring='accuracy', cv=skf
skfold_score_knnc = np.mean(skfold_scores_knnc)

skfold_scores_mlp = cross_val_score(clf_mlp_skfold, X, y, scoring='accuracy', cv=skf,
skfold_score_mlp = np.mean(skfold_scores_mlp)

## D Groupwise 10 fold

#defining classifiers
clf_tree_gkfold = tree.DecisionTreeClassifier(random_state=9)
clf_ranFor_gkfold = RandomForestClassifier(max_depth=2, random_state=9)
clf_knnc_gkfold =  KNeighborsClassifier(n_neighbors=3)
clf_mlp_gkfold = MLPClassifier(random_state=9, max_iter=300)

#defining the feature in df which have the grouping
groups_d = data['person']

#defining groupwise kfold object to define the evaulation strategy
gkf = GroupKFold(n_splits=10)

#Assigning the evaluation strategy to the classifier and then averging the mean of the
gkfold_scores_tree = cross_val_score(clf_tree_gkfold, X, y, scoring='accuracy', cv=gkf
gkfold_score_tree = np.mean(gkfold_scores_tree)

gkfold_scores_ranFor = cross_val_score(clf_ranFor_gkfold , X, y, scoring='accuracy', c
```

```
gkfold_score_ranFor = np.mean(gkfold_scores_ranFor)

gkfold_scores_knnc = cross_val_score(clf_knnc_gkfold, X, y, scoring='accuracy', cv=gkf
gkfold_score_knnc = np.mean(gkfold_scores_knnc)

gkfold_scores_mlp = cross_val_score(clf_mlp_gkfold, X, y, scoring='accuracy', cv=gkf,
gkfold_score_mlp = np.mean(gkfold_scores_mlp)


## E: Stratified groupwise 10-fold CV

#defining classifiers
clf_tree_sgkfold = tree.DecisionTreeClassifier(random_state=9)
clf_ranFor_sgkfold = RandomForestClassifier(max_depth=2, random_state=9)
clf_knnc_sgkfold =  KNeighborsClassifier(n_neighbors=3)
clf_mlp_sgkfold = MLPClassifier(random_state=9, max_iter=300)

#defining groupwise kfold object to define the evaulation strategy
sgkf = StratifiedGroupKFold(n_splits=10, random_state=9, shuffle=True)

#Assigning the evaluation strategy to the classifier and then averging the mean of the
sgkfold_scores_tree = cross_val_score(clf_tree_sgkfold, X, y, scoring='accuracy', cv=s
sgkfold_score_tree = np.mean(sgkfold_scores_tree)

sgkfold_scores_ranFor = cross_val_score(clf_ranFor_sgkfold , X, y, scoring='accuracy',
sgkfold_score_ranFor = np.mean(sgkfold_scores_ranFor)

sgkfold_scores_knnc = cross_val_score(clf_knnc_sgkfold, X, y, scoring='accuracy', cv=s
sgkfold_score_knnc = np.mean(sgkfold_scores_knnc)

sgkfold_scores_mlp = cross_val_score(clf_mlp_sgkfold, X, y, scoring='accuracy', cv=sgk
sgkfold_score_mlp = np.mean(sgkfold_scores_mlp)

## 5 - Full data

#defining classifiers
clf_dummy_f = DummyClassifier(strategy="stratified") #dummy clasifiers for the baselir
clf_tree_f = tree.DecisionTreeClassifier(random_state=9)
clf_ranFor_f = RandomForestClassifier(max_depth=2, random_state=9)
clf_knnc_f =  KNeighborsClassifier(n_neighbors=3)
clf_mlp_f = MLPClassifier(random_state=9, max_iter=300)

#seperating data into seperate train and test dataframes -

#Training Data
X_train_full = data[['G_front', 'G_vert', 'G_lat', 'ant_id', 'RSSI', 'phase', 'freq']]
y_train_full = data['activity']

#Test Data
X_test_full = test[['G_front', 'G_vert', 'G_lat', 'ant_id', 'RSSI', 'phase', 'freq']]
y_test_full = test['activity']

#Evaluating the data on the heldout data
clf_dummy_f.fit(X_train_full, y_train_full)
accuracy_5_base = clf_dummy_f.score(X_test_full, y_test_full)

clf_tree_f.fit(X_train_full, y_train_full)
accuracy_5_tree = clf_tree_f.score(X_test_full, y_test_full)

clf_ranFor_f.fit(X_train_full, y_train_full)
```

```
accuracy_5_ranFor = clf_ranFor_f.score(X_test_full, y_test_full)

clf_knnc_f.fit(X_train_full, y_train_full)
accuracy_5_knnc = clf_knnc_f.score(X_test_full, y_test_full)

clf_mlp_f.fit(X_train_full, y_train_full)
accuracy_5_mlp = clf_mlp_f.score(X_test_full, y_test_full)


#Tabulating for printing the data
table1 = [["Estimate","80-20 Split", "10-fold CV", "Stratified 10-fold CV", "Groupwise
          ["DT",accuracy_4_a_tree, kfold_score_tree, skfold_score_tree, gkfold_score_t
          ["RF", accuracy_4_a_ranFor, kfold_score_ranFor, skfold_score_ranFor, gkfold_
          ["3-NN", accuracy_4_a_knnc, kfold_score_knnc, skfold_score_knnc, gkfold_scor
          ["MLP", accuracy_4_a_mlp, kfold_score_mlp, skfold_score_mlp, gkfold_score_ml
          ]

table2 = [["Actual", "Heldout Accuracy"],
          ["Baseline", accuracy_5_base ],
          ["DT", accuracy_5_tree],
          ["RF", accuracy_5_ranFor],
          ["3-NN", accuracy_5_knnc],
          ["MLP",  accuracy_5_mlp]
          ]

#Calculating average of the errors found in each strategy per classifier
a = [accuracy_4_a_tree - accuracy_5_tree, accuracy_4_a_ranFor - accuracy_5_ranFor,  ac
avga = sum(a) / len(a)

b = [kfold_score_tree - accuracy_5_tree, kfold_score_ranFor- accuracy_5_ranFor,  kfold
avgb = np.mean(b)

c = [skfold_score_tree - accuracy_5_tree, skfold_score_ranFor- accuracy_5_ranFor, skfo
avgc = np.mean(c)

d = [ gkfold_score_tree - accuracy_5_tree, gkfold_score_ranFor- accuracy_5_ranFor, gkf
avgd = np.mean(d)

e = [sgkfold_score_tree - accuracy_5_tree, sgkfold_score_ranFor- accuracy_5_ranFor, sg
avge = np.mean(e)

table3 = [["Estimate","80-20 Split", "10-fold CV", "Stratified 10-fold CV", "Groupwise
          ["DT",accuracy_4_a_tree - accuracy_5_tree, kfold_score_tree - accuracy_5_tre
          ["RF", accuracy_4_a_ranFor - accuracy_5_ranFor, kfold_score_ranFor- accuracy
          ["3-NN", accuracy_4_a_knnc - accuracy_5_knnc, kfold_score_knnc - accuracy_5_
          ["MLP", accuracy_4_a_mlp - accuracy_5_mlp, kfold_score_mlp - accuracy_5_mlp,
          ["Avg", avga, avgb, avgc, avgd, avge]
          ]

#printing the tables

print("\n - Results from estimate by Evaluation Startegies -\n")
print(tabulate(table1, headers="firstrow", tablefmt="presto"))

print("\n - Results from actual Prediction on heldoutdata -")
print(tabulate(table2, headers="firstrow", tablefmt="presto"))

print("\n - Error report -")
print(tabulate(table3, headers="firstrow", tablefmt="presto"))
```

```
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
```

```
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
```

```
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
```

```
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
```

```
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\Users\katkh\anaconda3\lib\site-packages\sklearn\neighbors\_classification.py:228:
FutureWarning: Unlike other reduction functions (e.g. `skew`, `kurtosis`), the defaul
t behavior of `mode` typically preserves the axis it acts along. In SciPy 1.11.0, thi
s behavior will change: the default value of `keepdims` will become False, the `axis`
over which the statistic is taken will be eliminated, and the value None will no long
er be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
```

```
 - Results from estimate by Evaluation Startegies -

 Estimate   |   80-20 Split |   10-fold CV |   Stratified 10-fold CV |   Groupwise 10
-fold CV |   Stratified Groupwise 10-fold CV
------------+---------------+--------------+-------------------------+---------------
---------+-----------------------------------
 DT         |      0.987441 |     0.9802   |                0.986488 |
0.94564  |                           0.945684
 RF         |      0.949765 |     0.938081 |                0.938403 |
0.881577 |                           0.904673
 3-NN       |      0.971743 |     0.958506 |                0.958205 |
0.872893 |                           0.906864
 MLP        |      0.899529 |     0.889403 |                0.88937  |
0.809211 |                           0.851173

 - Results from actual Prediction on heldoutdata -
 Actual   |   Heldout Accuracy
----------+--------------------
 Baseline |           0.460583
 DT       |           0.712621
 RF       |           0.784294
 3-NN     |           0.672862
 MLP      |           0.802041

 - Error report -
 Estimate   |   80-20 Split |   10-fold CV |   Stratified 10-fold CV |   Groupwise 10
-fold CV |   Stratified Groupwise 10-fold CV
------------+---------------+--------------+-------------------------+---------------
---------+-----------------------------------
 DT         |     0.27482   |    0.267579  |                0.273867 |
0.233019  |                           0.233063
 RF         |     0.16547   |    0.153787  |                0.154109 |
0.0972823 |                           0.120379
 3-NN       |     0.298881  |    0.285645  |                0.285343 |
0.200031  |                           0.234002
 MLP        |     0.0974878 |    0.0873615 |                0.087329 |
0.0071697 |                           0.0491312
 Avg        |     0.209165  |    0.198593  |                0.200162 |
0.134376  |                           0.159144
```

In [ ]: