Untitled4

April 14, 2022

[1]:
```
import warnings
warnings.filterwarnings('ignore')
import numpy as np
import sklearn
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

[2]:
```
from sklearn.neighbors import KNeighborsClassifier
from sklearn.linear_model import LogisticRegression,SGDClassifier
from sklearn.svm import SVC, LinearSVC, NuSVC
from sklearn.tree import DecisionTreeClassifier
from sklearn.naive_bayes import GaussianNB,MultinomialNB
from sklearn.ensemble import RandomForestClassifier, AdaBoostClassifier,␣
 ↪GradientBoostingClassifier

from sklearn.metrics import classification_report
from sklearn.pipeline import Pipeline,make_pipeline
from sklearn.metrics import accuracy_score, log_loss
from sklearn.preprocessing import StandardScaler, MinMaxScaler
from sklearn.model_selection import train_test_split, GridSearchCV

from statsmodels.stats.outliers_influence import variance_inflation_factor
from sklearn.decomposition import PCA
from sklearn.feature_selection import (RFE, SelectKBest,
                                       SelectPercentile,RFECV)
```

[3]:
```
df = pd.read_csv('data/Train_Data.csv')
df.head()
```

[3]:
```
   battery_power  blue  clock_speed  dual_sim  fc  four_g  int_memory  m_dep  \
0            842     0          2.2         0   1       0           7    0.6
1           1021     1          0.5         1   0       1          53    0.7
2            563     1          0.5         1   2       1          41    0.9
3            615     1          2.5         0   0       0          10    0.8
4           1821     1          1.2         0  13       1          44    0.6

   mobile_wt  n_cores  pc  px_height  px_width   ram  sc_h  sc_w  talk_time  \
0        188        2   2         20       756  2549     9     7         19
1        136        3   6        905      1988  2631    17     3          7
2        145        5   6       1263      1716  2603    11     2          9
3        131        6   9       1216      1786  2769    16     8         11
4        141        2  14       1208      1212  1411     8     2         15

   three_g  touch_screen  wifi
0        0             0     1
1        1             1     0
2        1             1     0
3        1             0     0
4        1             1     0
```

[4]:
```
df2 = pd.read_csv('data/Traindata_classlabels.csv')
df2.head()
```

[4]:
```
   price_range
0            1
1            2
2            2
3            2
4            1
```

[5]:
```
data = {}
data_GS = {}
```

[6]:
```
def Classification(X_train, X_test, y_train, y_test,data,name:str,Gridsearch =␣
 ↪False):
    try:
        y_train = y_train.values.ravel()
        y_test = y_test.values.ravel()
    except:
        pass

    classifiers = [
    KNeighborsClassifier(),
    LogisticRegression()
    ]
    if Gridsearch ==True:
        clf_parameters = [
            {
                "clf__n_neighbors": np.arange(2,25 ,1),
                "clf__metric":␣
 ↪["cityblock","cosine","euclidean","l1","l2","manhattan","nan_euclidean",],
                "clf__weights": ["uniform", "distance"],
                "clf__algorithm": ["auto", "ball_tree", "kd_tree", "brute"],
        },
        {
                "clf__C": np.logspace(-2, -1.5, 30),
                "clf__penalty": ["l1", "l2", "elasticnet", "none"],
                "clf__solver": ["newton-cg", "lbfgs", "liblinear", "sag", "saga"],
        }
         ]
    else:
        clf_parameters = [{},{}]
    data[name] = {'Model':[],'Accuracy' :[],'f1_micro' :[],'f1_macro' :[]}
    dataint = {'Model':[],'Accuracy' :[],'f1_micro' :[],'f1_macro' :[]}

    i=1
    for classifier,clf_params in zip(classifiers,clf_parameters):
        pipe = Pipeline(steps=[('clf', classifier)])
        grid = GridSearchCV(pipe,clf_params,scoring='f1_macro',cv=10,n_jobs=-1)
        try:
            grid.fit(X_train, y_train)
            pred = grid.predict(X_test)
            print("_"*32)
            print(f'{i}.',classifier)
            print("_"*32)
            print(grid.best_params_)
            print(classification_report(y_test, pred))
            i+=1
            i1 = classifier.__class__.__name__
            i2 = sklearn.metrics.accuracy_score(y_test,pred)
            i3 = sklearn.metrics.f1_score(y_test,pred,average='micro')
            i4 = sklearn.metrics.f1_score(y_test,pred,average='macro')
            dataint['Model'].append(i1)
            dataint['Accuracy'].append(i2)
            dataint['f1_micro'].append(i3)
            dataint['f1_macro'].append(i4)
            print("-"*80)
            print("-"*80)

        except Exception as e: print(e)
    classifiers = [
        DecisionTreeClassifier(),
        SVC(),
        NuSVC(),
        RandomForestClassifier(),
        AdaBoostClassifier(),
        GradientBoostingClassifier(),
        SGDClassifier()
        ]
    base_estimators = classifiers
```

```
   if Gridsearch==True:
       clf_parameters = [
                         {
               'clf__criterion' : ["gini", "entropy"],
           'clf__max_features':['sqrt', 'log2',None],
       #             'max_depth':np.linspace(140,190,10),
                   'clf__ccp_alpha':np.logspace(-3,-2,20),#np.logspace(-2.
↪32,-2.3,20),
           "clf__max_leaf_nodes" : [None]+np.arange(30,40,5).tolist(),

           "clf__splitter" : ["best", "random"],
           "clf__min_samples_split":np.arange(2,50,10)
       },


           {
                   'clf__C':(0.8,2,30),
                   'clf__kernel':('linear','rbf','poly','sigmoid'),
                   'clf__decision_function_shape' : ['ovo', 'ovr'],
                   'clf__degree':np.arange(3,5,1)
                   },
           {
               'clf__nu': np.logspace(-1.15,-1.23,5),#np.logspace(-2,-1,30)
            'clf__kernel' : ['linear', 'poly', 'rbf', 'sigmoid' ]    ,
               'clf__gamma' : ['scale', 'auto'],
               'clf__degree':np.arange(3,5,1),
               'clf__decision_function_shape' : ['ovo', 'ovr'],
           },
           {
                'clf__n_estimators': [150,200,250],
           'clf__max_features': ['sqrt', 'log2',None],
           'clf__max_depth' : [None],#np.arange(4,15,2).tolist(),
           'clf__criterion' :['gini', 'entropy']    ,
               'clf__bootstrap' :[True],
       #         'clf__ccp_alpha':np.logspace(-2,1,10)
           },
           {
               'clf__base_estimator':
↪[RandomForestClassifier(),DecisionTreeClassifier(criterion='entropy'),SVC(),LogisticRegress
↪ 0.017433288221999882, 'penalty': 'l2', 'solver': 'newton-cg'})],
               'clf__algorithm' : ['SAMME', 'SAMME.R'],
               'clf__n_estimators': [50,100]
           },
           {
             'clf__loss' :['deviance', 'exponential'],
               'clf__criterion' : ['friedman_mse', 'squared_error'],
               'clf__max_features' : [ 'sqrt', 'log2'],
                'clf__learning_rate': np.logspace(-1,1,5),
                'clf__n_estimators':np.arange(100,1000,200)

            },
            {
                'clf__loss' :['hinge', 'log', 'modified_huber','squared_hinge',␣
↪'perceptron'],
                'clf__penalty' : ['l2', 'l1', 'elasticnet'],
                'clf__alpha' : np.logspace(-4,-2,30),
                'clf__learning_rate' :[␣
↪'constant','optimal','invscaling','adaptive'],
            }
            ]
    else:
        clf_parameters = [{}]*len(classifiers)
    for classifier,clf_params in zip(classifiers,clf_parameters):
        pipe = Pipeline(steps=[('standardscaler', StandardScaler()),('clf',␣
↪classifier)])
        grid = GridSearchCV(pipe,clf_params,scoring='f1_macro',cv=10,n_jobs=-1)
        try:
            grid.fit(X_train, y_train)
            pred = grid.predict(X_test)
            print("_"*32)
            print(f'{i}.',classifier)
            print("_"*32)
            print(grid.best_params_)
            print(classification_report(y_test, pred))
            i=i+1
            i1 = classifier.__class__.__name__
            i2 = sklearn.metrics.accuracy_score(y_test,pred)
            i3 = sklearn.metrics.f1_score(y_test,pred,average='micro')
            i4 = sklearn.metrics.f1_score(y_test,pred,average='macro')
            dataint['Model'].append(i1)
            dataint['Accuracy'].append(i2)
            dataint['f1_micro'].append(i3)
            dataint['f1_macro'].append(i4)
            print("-"*80)
            print("-"*80)

        except Exception as e: print(e)


    classifiers = [
    GaussianNB(),
    MultinomialNB()
    ]
    if Gridsearch==True:
        clf_parameters = [
            {
            'clf__var_smoothing':np.logspace(-20,-10,10)
            },
            {
                'clf__alpha':[0] + np.logspace(-2,5,5).tolist(),
                'clf__fit_prior':[True,False]
            }
             ]
    else:
        clf_parameters = [{}]*len(classifiers)

    for classifier,clf_params in zip(classifiers,clf_parameters):
        pipe = Pipeline(steps=[('minmaxscalar',MinMaxScaler()),('clf', classifier)])
        grid = GridSearchCV(pipe,clf_params,scoring='f1_macro',cv=10,n_jobs=-1)
        try:
            grid.fit(X_train, y_train)
            pred = grid.predict(X_test)
            print("_"*32)
            print(f'{i}.',classifier)
            print("_"*32)
            print(grid.best_params_)
            print(classification_report(y_test, pred))
            i=i+1
            i1 = classifier.__class__.__name__
            i2 = sklearn.metrics.accuracy_score(y_test,pred)
            i3 = sklearn.metrics.f1_score(y_test,pred,average='micro')
            i4 = sklearn.metrics.f1_score(y_test,pred,average='macro')
            dataint['Model'].append(i1)
            dataint['Accuracy'].append(i2)
            dataint['f1_micro'].append(i3)
            dataint['f1_macro'].append(i4)
            print("-"*80)
            print("-"*80)

        except Exception as e: print(e)

    data[name] = dataint
#     print(dataint)
```

1 With and Without Grid Search

1.1 Raw data

[7]:
```
X_train, X_test, y_train, y_test =␣
 ↪train_test_split(df,df2,random_state=5,shuffle=True,test_size=0.2)
Classification(X_train, X_test, y_train, y_test,data,'Raw')
print('WITH GridSearch'+'\n'+"-"*80+'\n'+"-"*80 )
Classification(X_train, X_test, y_train,␣
 ↪y_test,data_GS,'Raw_GS',Gridsearch=True)
```

```
--------------------------------
1. KNeighborsClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.98      0.97       103
           1       0.88      0.92      0.90        92
           2       0.88      0.87      0.88       101
           3       0.97      0.91      0.94       104

    accuracy                           0.92       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
2. LogisticRegression()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.88      0.86      0.87       103
           1       0.60      0.60      0.60        92
           2       0.49      0.46      0.47       101
           3       0.67      0.74      0.70       104

    accuracy                           0.67       400
   macro avg       0.66      0.66      0.66       400
weighted avg       0.66      0.67      0.67       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
3. DecisionTreeClassifier()

--------------------------------
```

```
{}
              precision    recall  f1-score   support

           0       0.93      0.90      0.92       103
           1       0.79      0.83      0.81        92
           2       0.81      0.80      0.81       101
           3       0.90      0.90      0.90       104

    accuracy                           0.86       400
   macro avg       0.86      0.86      0.86       400
weighted avg       0.86      0.86      0.86       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.94      0.95       103
           1       0.85      0.84      0.84        92
           2       0.80      0.89      0.85       101
           3       0.97      0.88      0.92       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
5. NuSVC()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.93      0.94       103
           1       0.83      0.83      0.83        92
           2       0.79      0.88      0.83       101
           3       0.97      0.88      0.92       104

    accuracy                           0.88       400
   macro avg       0.88      0.88      0.88       400
weighted avg       0.89      0.88      0.88       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

```
--------------------------------
6. RandomForestClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.98      0.97       103
           1       0.83      0.84      0.83        92
           2       0.82      0.80      0.81       101
           3       0.94      0.92      0.93       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.90      0.25      0.39       103
           1       0.46      0.86      0.60        92
           2       0.48      0.84      0.61       101
           3       0.96      0.21      0.35       104

    accuracy                           0.53       400
   macro avg       0.70      0.54      0.49       400
weighted avg       0.71      0.53      0.48       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.97      0.96       103
           1       0.85      0.88      0.87        92
           2       0.87      0.83      0.85       101
           3       0.94      0.93      0.94       104

    accuracy                           0.91       400
   macro avg       0.90      0.90      0.90       400
weighted avg       0.90      0.91      0.90       400
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.99      1.00      1.00       103
           1       0.52      0.58      0.55        92
           2       0.55      0.48      0.51       101
           3       0.96      0.99      0.98       104

    accuracy                           0.77       400
   macro avg       0.76      0.76      0.76       400
weighted avg       0.76      0.77      0.76       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.97      0.96       103
           1       0.71      0.71      0.71        92
           2       0.68      0.67      0.68       101
           3       0.91      0.90      0.91       104

    accuracy                           0.82       400
   macro avg       0.81      0.81      0.81       400
weighted avg       0.82      0.82      0.82       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.90      0.71      0.79       103
           1       0.46      0.60      0.52        92
           2       0.41      0.44      0.42       101
           3       0.66      0.59      0.62       104
    accuracy                           0.58       400
   macro avg       0.61      0.58      0.59       400
weighted avg       0.61      0.58      0.59       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
WITH GridSearch
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
1. KNeighborsClassifier()
--------------------------------
{'clf__algorithm': 'auto', 'clf__metric': 'cityblock', 'clf__n_neighbors': 18,
'clf__weights': 'distance'}
              precision    recall  f1-score   support

           0       0.97      0.98      0.98       103
           1       0.90      0.93      0.91        92
           2       0.90      0.90      0.90       101
           3       0.98      0.93      0.96       104

    accuracy                           0.94       400
   macro avg       0.94      0.94      0.94       400
weighted avg       0.94      0.94      0.94       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
2. LogisticRegression()
--------------------------------
{'clf__C': 0.02807216203941177, 'clf__penalty': 'l2', 'clf__solver': 'newton-
cg'}
              precision    recall  f1-score   support

           0       1.00      0.99      1.00       103
           1       0.98      0.99      0.98        92
           2       0.98      0.98      0.98       101
           3       0.99      0.99      0.99       104

    accuracy                           0.99       400
   macro avg       0.99      0.99      0.99       400
weighted avg       0.99      0.99      0.99       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
3. DecisionTreeClassifier()
--------------------------------
```

```
{'clf__ccp_alpha': 0.004832930238571752, 'clf__criterion': 'entropy',
'clf__max_features': None, 'clf__max_leaf_nodes': None,
'clf__min_samples_split': 2, 'clf__splitter': 'best'}
              precision    recall  f1-score   support

           0       0.90      0.92      0.91       103
           1       0.81      0.79      0.80        92
           2       0.82      0.85      0.83       101
           3       0.94      0.90      0.92       104

    accuracy                           0.87       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.87      0.87      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
4. SVC()
--------------------------------
{'clf__C': 30, 'clf__decision_function_shape': 'ovo', 'clf__degree': 3,
'clf__kernel': 'linear'}
              precision    recall  f1-score   support

           0       0.97      1.00      0.99       103
           1       0.99      0.96      0.97        92
           2       0.98      0.98      0.98       101
           3       0.99      0.99      0.99       104

    accuracy                           0.98       400
   macro avg       0.98      0.98      0.98       400
weighted avg       0.98      0.98      0.98       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
5. NuSVC()
--------------------------------
{'clf__decision_function_shape': 'ovo', 'clf__degree': 3, 'clf__gamma': 'scale',
'clf__kernel': 'linear', 'clf__nu': 0.0707945784384138}
              precision    recall  f1-score   support

           0       0.97      0.99      0.98       103
           1       0.96      0.96      0.96        92
           2       0.98      0.94      0.96       101
           3       0.97      0.99      0.98       104

    accuracy                           0.97       400
   macro avg       0.97      0.97      0.97       400
```

```
weighted avg       0.97      0.97      0.97       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
6. RandomForestClassifier()
--------------------------------
{'clf__bootstrap': True, 'clf__criterion': 'entropy', 'clf__max_depth': None,
'clf__max_features': None, 'clf__n_estimators': 150}
              precision    recall  f1-score   support

           0       0.95      0.96      0.96       103
           1       0.87      0.92      0.89        92
           2       0.92      0.87      0.89       101
           3       0.96      0.94      0.95       104

    accuracy                           0.93       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.93      0.93      0.93       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
7. AdaBoostClassifier()
--------------------------------
{'clf__algorithm': 'SAMME.R', 'clf__base_estimator': RandomForestClassifier(),
'clf__n_estimators': 100}
              precision    recall  f1-score   support

           0       0.95      0.98      0.97       103
           1       0.84      0.87      0.86        92
           2       0.83      0.80      0.81       101
           3       0.93      0.90      0.92       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{'clf__criterion': 'squared_error', 'clf__learning_rate': 0.1, 'clf__loss':
'deviance', 'clf__max_features': 'log2', 'clf__n_estimators': 300}
              precision    recall  f1-score   support

           0       0.98      0.97      0.98       103
           1       0.84      0.88      0.86        92
           2       0.83      0.84      0.84       101
           3       0.96      0.92      0.94       104

    accuracy                           0.91       400
   macro avg       0.90      0.90      0.90       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
9. SGDClassifier()
--------------------------------
{'clf__alpha': 0.004520353656360241, 'clf__learning_rate': 'optimal',
'clf__loss': 'log', 'clf__penalty': 'l1'}
              precision    recall  f1-score   support

           0       0.98      0.98      0.98       103
           1       0.68      0.79      0.73        92
           2       0.79      0.65      0.71       101
           3       0.98      0.99      0.99       104

    accuracy                           0.86       400
   macro avg       0.86      0.85      0.85       400
weighted avg       0.86      0.86      0.86       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
10. GaussianNB()
--------------------------------
{'clf__var_smoothing': 1e-20}
              precision    recall  f1-score   support

           0       0.95      0.97      0.96       103
           1       0.71      0.71      0.71        92
           2       0.68      0.67      0.68       101
           3       0.91      0.90      0.91       104

    accuracy                           0.82       400
   macro avg       0.81      0.81      0.81       400
weighted avg       0.82      0.82      0.82       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
11. MultinomialNB()
--------------------------------
```

```
{'clf__alpha': 0, 'clf__fit_prior': False}
              precision    recall  f1-score   support

           0       0.88      0.78      0.82       103
           1       0.49      0.50      0.49        92
           2       0.41      0.46      0.43       101
           3       0.63      0.62      0.63       104

    accuracy                           0.59       400
   macro avg       0.60      0.59      0.59       400
weighted avg       0.61      0.59      0.60       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

1.2 Feature Engineering

[8]:

```
# df4.drop('four_g',axis = 1,inplace=True)
df4 = df.copy()
df4['sc_diag'] = np.sqrt(df['sc_h']**2 + df['sc_w']**2)
df4['sc_area'] = df['sc_h'] * df['sc_w']
df4['px_diag'] = np.sqrt(df['px_height']**2 + df['px_width']**2)
df4['px_area'] = df['px_width'] * df['px_height']
df4['talk_per_mAh'] = df['talk_time']/ df['battery_power']
df4['connectivity'] = df['three_g'] + df['four_g'] + df['blue'] + df['wifi'] +␣
 ↪df['dual_sim']
df4['ram_per_core'] = df['ram'] / df['n_cores']
df4['weight_by_thickness'] = df['mobile_wt'] / df['m_dep']
df4['total_clock_speed'] = df['clock_speed'] * df['n_cores']
df4['fc_res'] = [1 if df['fc'][i] >=8 else 0 for i in range(len(df)) ]
df4['pc_res'] = [1 if df['pc'][i] >=8 else 0 for i in range(len(df)) ]
df4['sc_res'] = [1 if( df['px_width'][i] * df['px_height'][i] ) >= 921600 else␣
 ↪0 for i in range(len(df)) ]
df4['DPIx'] = df['px_width'] / (df['sc_w']*0.394 + 1)
df4['DPIy'] = df['px_height'] / (df['sc_h']*0.394 + 1)
df4['DPI'] = (df['px_width'] * df['px_height']) / (df4['sc_area'] * (0.394)**2␣
 ↪+ 1)
df4
X_train_e, X_test_e, y_train_e, y_test_e =␣
 ↪train_test_split(df4,df2,random_state=5,shuffle=True,test_size=0.2)
```

1.3 After Feature Engineering

[9]:

```
Classification(X_train_e, X_test_e, y_train_e, y_test_e,data,'Feature_eng')
print('WITH GridSearch'+'\n'+"-"*80+'\n'+"-"*80 )
Classification(X_train_e, X_test_e, y_train_e,␣
 ↪y_test_e,data_GS,'Feature_eng_GS',Gridsearch=True)
```

```
--------------------------------
1. KNeighborsClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.33      0.45      0.38       103
           1       0.26      0.27      0.27        92
           2       0.24      0.23      0.24       101
           3       0.29      0.20      0.24       104

    accuracy                           0.29       400
   macro avg       0.28      0.29      0.28       400
weighted avg       0.28      0.29      0.28       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
2. LogisticRegression()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.92      0.68      0.78       103
           1       0.42      0.58      0.48        92
           2       0.32      0.24      0.27       101
           3       0.61      0.71      0.65       104

    accuracy                           0.55       400
   macro avg       0.57      0.55      0.55       400
weighted avg       0.57      0.55      0.55       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.92      0.94       103
           1       0.84      0.87      0.86        92
           2       0.82      0.83      0.83       101
           3       0.90      0.90      0.90       104

    accuracy                           0.88       400
   macro avg       0.88      0.88      0.88       400
weighted avg       0.88      0.88      0.88       400
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
4. SVC()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.94      0.95       103
           1       0.80      0.82      0.81        92
           2       0.77      0.81      0.79       101
           3       0.94      0.88      0.91       104

    accuracy                           0.86       400
   macro avg       0.87      0.86      0.86       400
weighted avg       0.87      0.86      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
5. NuSVC()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.95      0.96       103
           1       0.81      0.82      0.81        92
           2       0.77      0.81      0.79       101
           3       0.94      0.89      0.92       104

    accuracy                           0.87       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.87      0.87      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
6. RandomForestClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.97      0.97       103
           1       0.82      0.82      0.82        92
           2       0.81      0.84      0.83       101
           3       0.97      0.93      0.95       104
    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.93      0.24      0.38       103
           1       0.45      0.84      0.58        92
           2       0.52      0.81      0.63       101
           3       0.95      0.39      0.56       104

    accuracy                           0.56       400
   macro avg       0.71      0.57      0.54       400
weighted avg       0.72      0.56      0.54       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.98      0.97      0.98       103
           1       0.84      0.91      0.87        92
           2       0.89      0.83      0.86       101
           3       0.96      0.96      0.96       104

    accuracy                           0.92       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       103
           1       0.49      0.36      0.42        92
           2       0.52      0.65      0.58       101
           3       0.99      0.97      0.98       104

    accuracy                           0.76       400
   macro avg       0.75      0.75      0.74       400
weighted avg       0.76      0.76      0.75       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.89      0.94      0.92       103
           1       0.64      0.60      0.62        92
           2       0.60      0.68      0.64       101
           3       0.92      0.80      0.86       104

    accuracy                           0.76       400
   macro avg       0.76      0.76      0.76       400
weighted avg       0.77      0.76      0.76       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.75      0.71      0.73       103
           1       0.37      0.39      0.38        92
           2       0.44      0.48      0.46       101
           3       0.56      0.53      0.54       104

    accuracy                           0.53       400
   macro avg       0.53      0.53      0.53       400
weighted avg       0.54      0.53      0.53       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
WITH GridSearch
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
1. KNeighborsClassifier()
```

```
--------------------------------
{'clf__algorithm': 'auto', 'clf__metric': 'cosine', 'clf__n_neighbors': 15,
'clf__weights': 'distance'}
              precision    recall  f1-score   support

           0       0.67      0.50      0.57       103
           1       0.30      0.32      0.31        92
           2       0.37      0.41      0.38       101
           3       0.42      0.46      0.44       104

    accuracy                           0.42       400
   macro avg       0.44      0.42      0.43       400
weighted avg       0.44      0.42      0.43       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
2. LogisticRegression()
--------------------------------
{'clf__C': 0.01, 'clf__penalty': 'none', 'clf__solver': 'newton-cg'}
              precision    recall  f1-score   support

           0       0.91      0.92      0.92       103
           1       0.73      0.75      0.74        92
           2       0.67      0.63      0.65       101
           3       0.81      0.82      0.81       104

    accuracy                           0.78       400
   macro avg       0.78      0.78      0.78       400
weighted avg       0.78      0.78      0.78       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{'clf__ccp_alpha': 0.00206913808111479, 'clf__criterion': 'entropy',
'clf__max_features': None, 'clf__max_leaf_nodes': None,
'clf__min_samples_split': 2, 'clf__splitter': 'best'}
              precision    recall  f1-score   support

           0       0.95      0.91      0.93       103
           1       0.82      0.89      0.85        92
           2       0.85      0.85      0.85       101
           3       0.94      0.90      0.92       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
```

```
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
4. SVC()
--------------------------------
{'clf__C': 30, 'clf__decision_function_shape': 'ovo', 'clf__degree': 3,
'clf__kernel': 'linear'}
              precision    recall  f1-score   support

           0       0.98      0.99      0.99       103
           1       0.96      0.97      0.96        92
           2       0.96      0.96      0.96       101
           3       0.99      0.97      0.98       104

    accuracy                           0.97       400
   macro avg       0.97      0.97      0.97       400
weighted avg       0.97      0.97      0.97       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
5. NuSVC()
--------------------------------
{'clf__decision_function_shape': 'ovo', 'clf__degree': 3, 'clf__gamma': 'scale',
'clf__kernel': 'linear', 'clf__nu': 0.06165950018614822}
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       103
           1       0.97      0.97      0.97        92
           2       0.99      0.95      0.97       101
           3       0.98      1.00      0.99       104

    accuracy                           0.98       400
   macro avg       0.98      0.98      0.98       400
weighted avg       0.98      0.98      0.98       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
6. RandomForestClassifier()
--------------------------------
{'clf__bootstrap': True, 'clf__criterion': 'entropy', 'clf__max_depth': None,
'clf__max_features': None, 'clf__n_estimators': 250}
              precision    recall  f1-score   support

           0       0.96      0.94      0.95       103
           1       0.84      0.95      0.89        92
           2       0.95      0.89      0.92       101
           3       0.99      0.96      0.98       104

    accuracy                           0.94       400
   macro avg       0.94      0.94      0.93       400
weighted avg       0.94      0.94      0.94       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()
--------------------------------
{'clf__algorithm': 'SAMME.R', 'clf__base_estimator': RandomForestClassifier(),
'clf__n_estimators': 100}
              precision    recall  f1-score   support

           0       0.98      0.91      0.94       103
           1       0.76      0.85      0.80        92
           2       0.80      0.78      0.79       101
           3       0.94      0.92      0.93       104

    accuracy                           0.87       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.87      0.87      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{'clf__criterion': 'friedman_mse', 'clf__learning_rate': 1.0, 'clf__loss':
'deviance', 'clf__max_features': 'sqrt', 'clf__n_estimators': 300}
              precision    recall  f1-score   support

           0       0.99      0.94      0.97       103
           1       0.83      0.87      0.85        92
           2       0.82      0.83      0.82       101
           3       0.93      0.92      0.93       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()
```

```
--------------------------------
{'clf__alpha': 0.01, 'clf__learning_rate': 'optimal', 'clf__loss': 'log',
'clf__penalty': 'l1'}
              precision    recall  f1-score   support

           0       0.97      0.98      0.98       103
           1       0.73      0.75      0.74        92
           2       0.77      0.73      0.75       101
           3       0.96      0.98      0.97       104

    accuracy                           0.86       400
   macro avg       0.86      0.86      0.86       400
weighted avg       0.86      0.86      0.86       400

---------------------------------------------------------------------------------
---------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()

--------------------------------
{'clf__var_smoothing': 1e-20}
              precision    recall  f1-score   support

           0       0.89      0.94      0.92       103
           1       0.64      0.60      0.62        92
           2       0.60      0.68      0.64       101
           3       0.92      0.80      0.86       104

    accuracy                           0.76       400
   macro avg       0.76      0.76      0.76       400
weighted avg       0.77      0.76      0.76       400

---------------------------------------------------------------------------------
---------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()

--------------------------------
{'clf__alpha': 0, 'clf__fit_prior': True}
              precision    recall  f1-score   support

           0       0.75      0.71      0.73       103
           1       0.37      0.39      0.38        92
           2       0.44      0.48      0.46       101
           3       0.56      0.53      0.54       104

    accuracy                           0.53       400
   macro avg       0.53      0.53      0.53       400
weighted avg       0.54      0.53      0.53       400
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

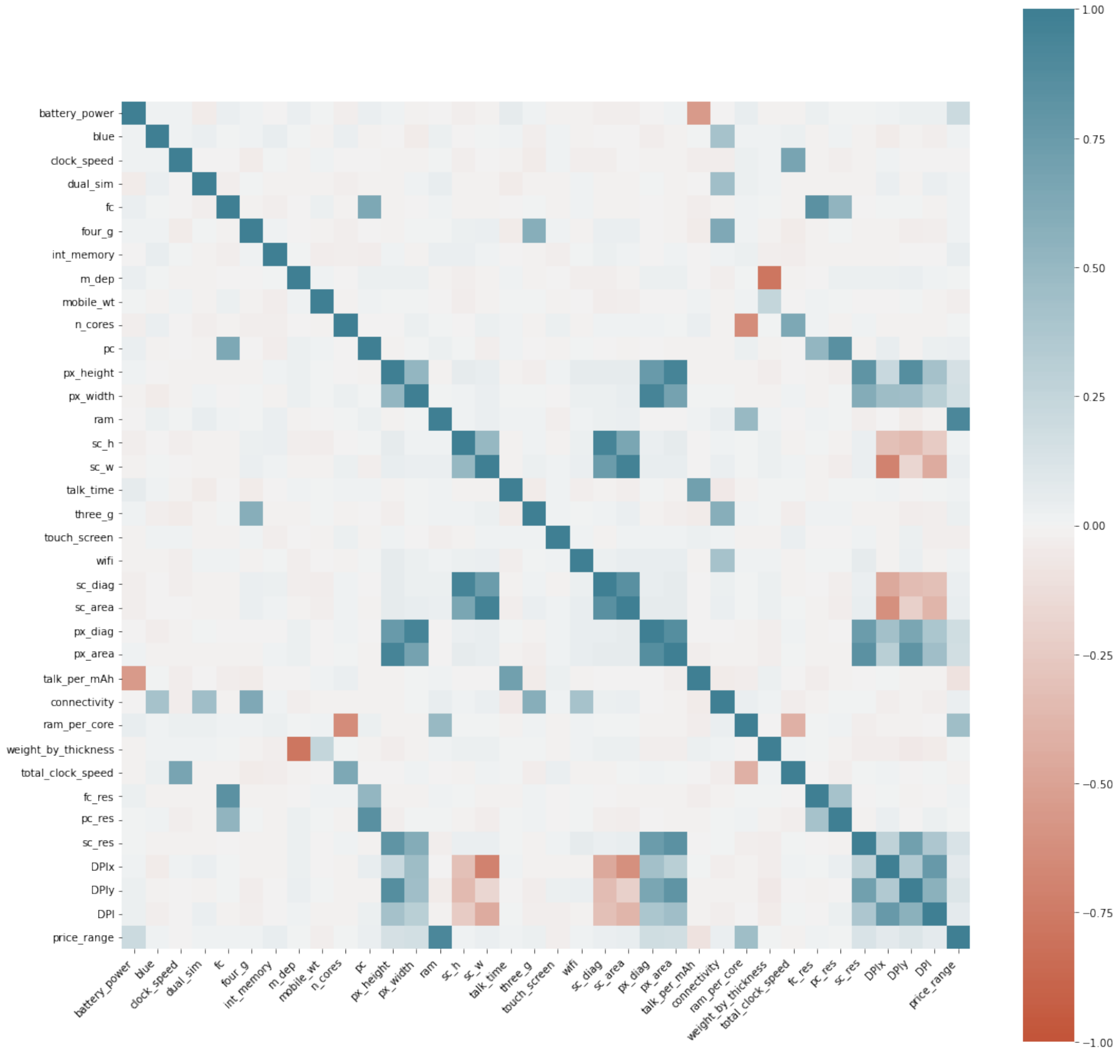
1.4 Feature Selection

1.4.1 Correlation

[10]:

```
df5 = df4.copy()
df5['price_range'] = df2
corr = df5.corr()
plt.figure(figsize = (18,18))
ax = sns.heatmap(
    corr,
    vmin=-1, vmax=1, center=0,
    cmap=sns.diverging_palette(20, 220, n=200),
    square=True
)
ax.set_xticklabels(
    ax.get_xticklabels(),
    rotation=45,
    horizontalalignment='right'
);
d = 0.01
col = []
for idx,i in enumerate(corr['price_range']):
    if i <= d and i >= -d:
        col.append(corr.columns[idx])
        print(corr.columns[idx])
df5 = df5.drop(col,axis = 1)
df5 = df5.drop('price_range',axis = 1)
df5
X_train_ec, X_test_ec, y_train_ec, y_test_ec =␣
 ↪train_test_split(df5,df2,random_state=5,shuffle=True,test_size=0.2)
```

```
clock_speed
m_dep
n_cores
weight_by_thickness
total_clock_speed
pc_res
```



1.4.2 After Feature Engg + Correlation

```
[11]: Classification(X_train_ec, X_test_ec, y_train_ec, y_test_ec,data,'Feature_engg␣
       ↪+ Corr')
      print('WITH GridSearch'+'\n'+"-"*80+'\n'+"-"*80 )
      Classification(X_train_ec, X_test_ec, y_train_ec,␣
       ↪y_test_ec,data_GS,'Feature_engg + Corr_GS',Gridsearch=True)
```

```
--------------------------------
1. KNeighborsClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.33      0.45      0.38       103
           1       0.26      0.27      0.27        92
           2       0.25      0.24      0.24       101
           3       0.30      0.20      0.24       104

    accuracy                           0.29       400
   macro avg       0.29      0.29      0.28       400
weighted avg       0.29      0.29      0.28       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
2. LogisticRegression()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.92      0.67      0.78       103
           1       0.40      0.63      0.49        92
           2       0.35      0.14      0.20       101
           3       0.56      0.75      0.64       104

    accuracy                           0.55       400
   macro avg       0.56      0.55      0.53       400
weighted avg       0.56      0.55      0.53       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.93      0.93      0.93       103
           1       0.85      0.86      0.85        92
           2       0.85      0.85      0.85       101
           3       0.92      0.91      0.92       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
4. SVC()
--------------------------------
{}
```

```
              precision    recall  f1-score   support

           0       0.98      0.94      0.96       103
           1       0.82      0.87      0.84        92
           2       0.81      0.82      0.81       101
           3       0.94      0.90      0.92       104

    accuracy                           0.89       400
   macro avg       0.89      0.88      0.88       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
5. NuSVC()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.98      0.94      0.96       103
           1       0.81      0.87      0.84        92
           2       0.80      0.83      0.82       101
           3       0.96      0.89      0.93       104

    accuracy                           0.89       400
   macro avg       0.89      0.88      0.88       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
6. RandomForestClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.99      0.95      0.97       103
           1       0.79      0.88      0.84        92
           2       0.81      0.84      0.83       101
           3       1.00      0.90      0.95       104

    accuracy                           0.90       400
   macro avg       0.90      0.89      0.90       400
weighted avg       0.90      0.90      0.90       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
```

```
7. AdaBoostClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.87      0.46      0.60       103
           1       0.51      0.77      0.61        92
           2       0.67      0.72      0.70       101
           3       0.84      0.79      0.81       104

    accuracy                           0.68       400
   macro avg       0.72      0.68      0.68       400
weighted avg       0.73      0.68      0.68       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.98      0.97      0.98       103
           1       0.85      0.89      0.87        92
           2       0.86      0.82      0.84       101
           3       0.94      0.94      0.94       104

    accuracy                           0.91       400
   macro avg       0.91      0.91      0.91       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
9. SGDClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       1.00      0.97      0.99       103
           1       0.58      0.57      0.57        92
           2       0.61      0.64      0.63       101
           3       0.98      0.98      0.98       104

    accuracy                           0.80       400
   macro avg       0.79      0.79      0.79       400
weighted avg       0.80      0.80      0.80       400
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
10. GaussianNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.90      0.93      0.91       103
           1       0.64      0.63      0.64        92
           2       0.61      0.69      0.65       101
           3       0.93      0.80      0.86       104

    accuracy                           0.77       400
   macro avg       0.77      0.76      0.77       400
weighted avg       0.78      0.77      0.77       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
11. MultinomialNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.75      0.70      0.72       103
           1       0.41      0.42      0.42        92
           2       0.41      0.48      0.44       101
           3       0.58      0.51      0.54       104

    accuracy                           0.53       400
   macro avg       0.54      0.53      0.53       400
weighted avg       0.54      0.53      0.53       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
WITH GridSearch
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
1. KNeighborsClassifier()
--------------------------------
{'clf__algorithm': 'auto', 'clf__metric': 'cosine', 'clf__n_neighbors': 18,
'clf__weights': 'distance'}
              precision    recall  f1-score   support

           0       0.71      0.52      0.60       103
           1       0.30      0.32      0.31        92
           2       0.38      0.42      0.40       101
           3       0.42      0.47      0.45       104

    accuracy                           0.43       400
   macro avg       0.45      0.43      0.44       400
weighted avg       0.46      0.43      0.44       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
2. LogisticRegression()
--------------------------------
{'clf__C': 0.029209037170322485, 'clf__penalty': 'l2', 'clf__solver': 'newton-
cg'}
              precision    recall  f1-score   support

           0       0.90      0.92      0.91       103
           1       0.75      0.72      0.73        92
           2       0.65      0.66      0.66       101
           3       0.80      0.79      0.79       104

    accuracy                           0.78       400
   macro avg       0.77      0.77      0.77       400
weighted avg       0.77      0.78      0.77       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{'clf__ccp_alpha': 0.00379269019073225, 'clf__criterion': 'entropy',
'clf__max_features': None, 'clf__max_leaf_nodes': None,
'clf__min_samples_split': 2, 'clf__splitter': 'best'}
              precision    recall  f1-score   support

           0       0.94      0.93      0.94       103
           1       0.82      0.90      0.86        92
           2       0.88      0.82      0.85       101
           3       0.93      0.92      0.93       104

    accuracy                           0.90       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.90      0.90      0.90       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()
```

```
--------------------------------
{'clf__C': 30, 'clf__decision_function_shape': 'ovo', 'clf__degree': 3,
'clf__kernel': 'linear'}
              precision    recall  f1-score   support

           0       0.99      0.99      0.99       103
           1       0.97      0.98      0.97        92
           2       0.96      0.97      0.97       101
           3       0.99      0.97      0.98       104

    accuracy                           0.98       400
   macro avg       0.98      0.98      0.98       400
weighted avg       0.98      0.98      0.98       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
5. NuSVC()
--------------------------------
{'clf__decision_function_shape': 'ovo', 'clf__degree': 3, 'clf__gamma': 'scale',
'clf__kernel': 'linear', 'clf__nu': 0.06165950018614822}
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       103
           1       0.98      0.97      0.97        92
           2       0.98      0.96      0.97       101
           3       0.98      0.99      0.99       104

    accuracy                           0.98       400
   macro avg       0.98      0.98      0.98       400
weighted avg       0.98      0.98      0.98       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
6. RandomForestClassifier()
--------------------------------
{'clf__bootstrap': True, 'clf__criterion': 'entropy', 'clf__max_depth': None,
'clf__max_features': None, 'clf__n_estimators': 200}
              precision    recall  f1-score   support

           0       0.96      0.95      0.96       103
           1       0.87      0.92      0.89        92
           2       0.91      0.89      0.90       101
           3       0.97      0.94      0.96       104

    accuracy                           0.93       400
   macro avg       0.93      0.93      0.93       400
```

```
weighted avg       0.93      0.93      0.93       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
7. AdaBoostClassifier()
--------------------------------
{'clf__algorithm': 'SAMME.R', 'clf__base_estimator': RandomForestClassifier(),
'clf__n_estimators': 100}
              precision    recall  f1-score   support

           0       0.98      0.95      0.97       103
           1       0.82      0.87      0.85        92
           2       0.84      0.86      0.85       101
           3       0.98      0.93      0.96       104

    accuracy                           0.91       400
   macro avg       0.91      0.90      0.90       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{'clf__criterion': 'friedman_mse', 'clf__learning_rate': 0.1, 'clf__loss':
'deviance', 'clf__max_features': 'sqrt', 'clf__n_estimators': 700}
              precision    recall  f1-score   support

           0       0.98      0.96      0.97       103
           1       0.84      0.85      0.84        92
           2       0.81      0.83      0.82       101
           3       0.94      0.92      0.93       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
9. SGDClassifier()
--------------------------------
{'clf__alpha': 0.007278953843983146, 'clf__learning_rate': 'optimal',
'clf__loss': 'log', 'clf__penalty': 'l1'}
              precision    recall  f1-score   support

           0       0.98      0.98      0.98       103
           1       0.73      0.76      0.74        92
           2       0.78      0.75      0.77       101
           3       0.99      0.99      0.99       104

    accuracy                           0.88       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.88      0.88      0.88       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
10. GaussianNB()
--------------------------------
{'clf__var_smoothing': 1e-20}
              precision    recall  f1-score   support

           0       0.90      0.93      0.91       103
           1       0.64      0.63      0.64        92
           2       0.61      0.69      0.65       101
           3       0.93      0.80      0.86       104

    accuracy                           0.77       400
   macro avg       0.77      0.76      0.77       400
weighted avg       0.78      0.77      0.77       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
11. MultinomialNB()
--------------------------------
{'clf__alpha': 0, 'clf__fit_prior': True}
              precision    recall  f1-score   support

           0       0.76      0.70      0.73       103
           1       0.41      0.42      0.41        92
           2       0.41      0.48      0.44       101
           3       0.58      0.51      0.54       104

    accuracy                           0.53       400
   macro avg       0.54      0.53      0.53       400
weighted avg       0.54      0.53      0.53       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

1.4.3 Variance Inflation Factor

[12]:

```
scalar = StandardScaler()
X_tranform = scalar.fit_transform(df5)
df6 = pd.DataFrame(X_tranform,columns=df5.columns)
df6
vif = [variance_inflation_factor(df6.values,i) for i in range(df6.shape[1])]
d = 5 # Threshold for Variance Inflation,
col = []
for idx,i in enumerate(vif):
    if i >= d:
        col.append(df6.columns[idx])
        print(df6.columns[idx])
df6  = df6.drop(col,axis = 1)
X_train_ecv, X_test_ecv, y_train_ecv, y_test_ecv =␣
 ↪train_test_split(df6,df2,random_state=5,shuffle=True,test_size=0.2)
```

```
blue
dual_sim
four_g
px_height
px_width
sc_h
sc_w
talk_time
three_g
wifi
sc_diag
sc_area
px_diag
px_area
talk_per_mAh
connectivity
DPIx
DPIy
```

[13]:

```
Classification(X_train_ecv, X_test_ecv, y_train_ecv,␣
 ↪y_test_ecv,data,'Engg+corr+VIF')
print('WITH GridSearch'+'\n'+"-"*80+'\n'+"-"*80 )
Classification(X_train_ecv, X_test_ecv, y_train_ecv,␣
 ↪y_test_ecv,data_GS,'Engg+corr+VIF_GS',Gridsearch=True)
```

```
--------------------------------
1. KNeighborsClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.82      0.84      0.83       103
           1       0.57      0.66      0.61        92
           2       0.59      0.57      0.58       101
           3       0.84      0.71      0.77       104

    accuracy                           0.70       400
   macro avg       0.70      0.70      0.70       400
weighted avg       0.71      0.70      0.70       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
2. LogisticRegression()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.93      0.97      0.95       103
           1       0.85      0.84      0.84        92
           2       0.84      0.83      0.84       101
           3       0.94      0.92      0.93       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.93      0.89      0.91       103
           1       0.80      0.82      0.81        92
           2       0.78      0.79      0.79       101
           3       0.89      0.89      0.89       104

    accuracy                           0.85       400
   macro avg       0.85      0.85      0.85       400
weighted avg       0.85      0.85      0.85       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()
--------------------------------
{}
```

```
              precision    recall  f1-score   support

           0       0.94      0.94      0.94       103
           1       0.84      0.84      0.84        92
           2       0.79      0.81      0.80       101
           3       0.90      0.88      0.89       104

    accuracy                           0.87       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.87      0.87      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
5. NuSVC()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.93      0.94       103
           1       0.83      0.84      0.83        92
           2       0.79      0.80      0.79       101
           3       0.89      0.88      0.89       104

    accuracy                           0.86       400
   macro avg       0.86      0.86      0.86       400
weighted avg       0.87      0.86      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
6. RandomForestClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.94      0.95       103
           1       0.78      0.87      0.82        92
           2       0.82      0.75      0.78       101
           3       0.91      0.91      0.91       104

    accuracy                           0.87       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.87      0.87      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
```

```
7. AdaBoostClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.94      0.58      0.72       103
           1       0.58      0.90      0.71        92
           2       0.64      0.70      0.67       101
           3       0.84      0.66      0.74       104

    accuracy                           0.71       400
   macro avg       0.75      0.71      0.71       400
weighted avg       0.76      0.71      0.71       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.94      0.95       103
           1       0.79      0.84      0.81        92
           2       0.86      0.76      0.81       101
           3       0.91      0.97      0.94       104

    accuracy                           0.88       400
   macro avg       0.88      0.88      0.88       400
weighted avg       0.88      0.88      0.88       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.94      0.98      0.96       103
           1       0.57      0.70      0.62        92
           2       0.65      0.46      0.53       101
           3       0.91      0.95      0.93       104

    accuracy                           0.78       400
   macro avg       0.77      0.77      0.76       400
weighted avg       0.77      0.78      0.77       400
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.91      0.93      0.92       103
           1       0.68      0.67      0.68        92
           2       0.63      0.63      0.63       101
           3       0.86      0.85      0.85       104

    accuracy                           0.78       400
   macro avg       0.77      0.77      0.77       400
weighted avg       0.77      0.78      0.77       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.85      0.79      0.82       103
           1       0.40      0.48      0.43        92
           2       0.42      0.32      0.36       101
           3       0.56      0.63      0.60       104

    accuracy                           0.56       400
   macro avg       0.56      0.55      0.55       400
weighted avg       0.56      0.56      0.56       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
WITH GridSearch
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
1. KNeighborsClassifier()

--------------------------------
{'clf__algorithm': 'auto', 'clf__metric': 'cityblock', 'clf__n_neighbors': 24,
'clf__weights': 'distance'}
              precision    recall  f1-score   support

           0       0.86      0.88      0.87       103
           1       0.70      0.71      0.70        92
           2       0.70      0.73      0.71       101
           3       0.88      0.81      0.84       104

    accuracy                           0.79       400
   macro avg       0.78      0.78      0.78       400
weighted avg       0.79      0.79      0.79       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
2. LogisticRegression()

--------------------------------
{'clf__C': 0.01, 'clf__penalty': 'none', 'clf__solver': 'saga'}
              precision    recall  f1-score   support

           0       0.93      0.97      0.95       103
           1       0.84      0.84      0.84        92
           2       0.85      0.82      0.83       101
           3       0.94      0.93      0.94       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
3. DecisionTreeClassifier()

--------------------------------
{'clf__ccp_alpha': 0.0011288378916846896, 'clf__criterion': 'gini',
'clf__max_features': None, 'clf__max_leaf_nodes': None,
'clf__min_samples_split': 12, 'clf__splitter': 'random'}
              precision    recall  f1-score   support

           0       0.93      0.97      0.95       103
           1       0.84      0.86      0.85        92
           2       0.87      0.70      0.78       101
           3       0.84      0.94      0.89       104

    accuracy                           0.87       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.87      0.87      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()

--------------------------------
```

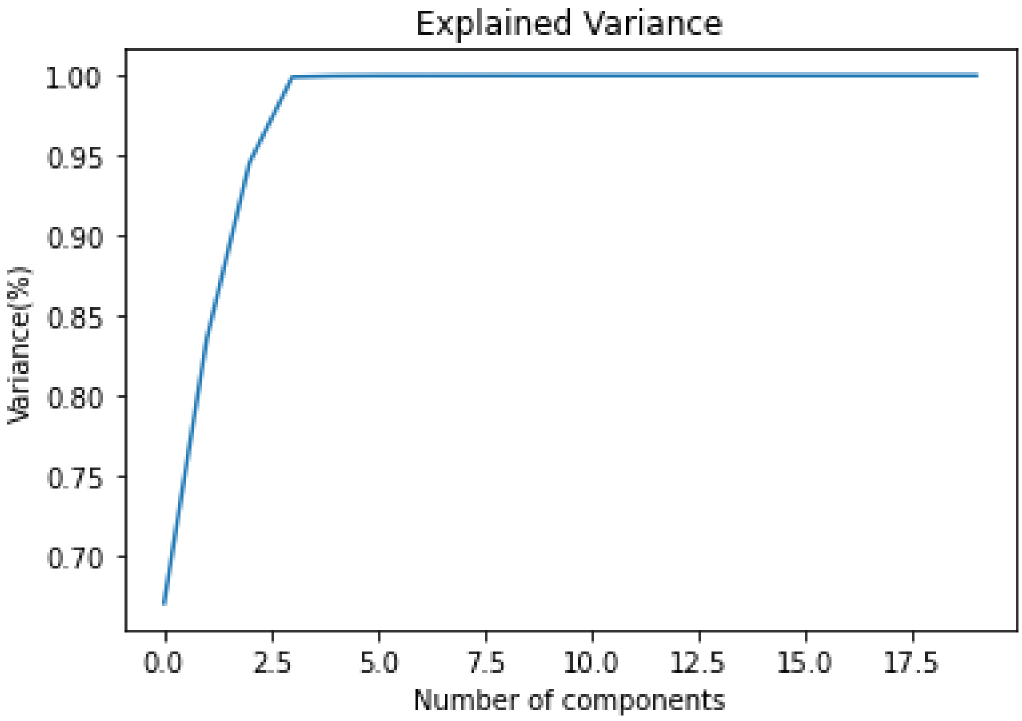
```
{'clf__C': 0.8, 'clf__decision_function_shape': 'ovo', 'clf__degree': 3,
'clf__kernel': 'linear'}
              precision    recall  f1-score   support

           0       0.93      0.97      0.95       103
           1       0.84      0.84      0.84        92
           2       0.83      0.81      0.82       101
           3       0.93      0.91      0.92       104

    accuracy                           0.89       400
   macro avg       0.88      0.88      0.88       400
weighted avg       0.88      0.89      0.88       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
5. NuSVC()
--------------------------------
{'clf__decision_function_shape': 'ovo', 'clf__degree': 3, 'clf__gamma': 'scale',
'clf__kernel': 'linear', 'clf__nu': 0.06456542290346556}
              precision    recall  f1-score   support

           0       0.91      0.93      0.92       103
           1       0.76      0.74      0.75        92
           2       0.72      0.75      0.74       101
           3       0.90      0.87      0.88       104

    accuracy                           0.82       400
   macro avg       0.82      0.82      0.82       400
weighted avg       0.83      0.82      0.83       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
6. RandomForestClassifier()
--------------------------------
{'clf__bootstrap': True, 'clf__criterion': 'entropy', 'clf__max_depth': None,
'clf__max_features': None, 'clf__n_estimators': 250}
              precision    recall  f1-score   support

           0       0.95      0.94      0.95       103
           1       0.80      0.90      0.85        92
           2       0.90      0.77      0.83       101
           3       0.93      0.95      0.94       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.90      0.89      0.89       400
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()
--------------------------------
{'clf__algorithm': 'SAMME', 'clf__base_estimator': RandomForestClassifier(),
'clf__n_estimators': 50}
              precision    recall  f1-score   support

           0       0.96      0.93      0.95       103
           1       0.79      0.88      0.83        92
           2       0.84      0.76      0.80       101
           3       0.91      0.92      0.92       104

    accuracy                           0.88       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.88      0.88      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{'clf__criterion': 'friedman_mse', 'clf__learning_rate': 0.1, 'clf__loss':
'deviance', 'clf__max_features': 'sqrt', 'clf__n_estimators': 700}
              precision    recall  f1-score   support

           0       0.97      0.96      0.97       103
           1       0.82      0.88      0.85        92
           2       0.85      0.78      0.81       101
           3       0.92      0.94      0.93       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()
--------------------------------
{'clf__alpha': 0.005298316906283708, 'clf__learning_rate': 'optimal',
'clf__loss': 'log', 'clf__penalty': 'l1'}
              precision    recall  f1-score   support

           0       0.92      0.97      0.94       103
           1       0.71      0.66      0.69        92
           2       0.72      0.70      0.71       101
           3       0.92      0.94      0.93       104

    accuracy                           0.82       400
   macro avg       0.82      0.82      0.82       400
weighted avg       0.82      0.82      0.82       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()
--------------------------------
{'clf__var_smoothing': 1e-20}
              precision    recall  f1-score   support

           0       0.91      0.93      0.92       103
           1       0.68      0.67      0.68        92
           2       0.63      0.63      0.63       101
           3       0.86      0.85      0.85       104

    accuracy                           0.78       400
   macro avg       0.77      0.77      0.77       400
weighted avg       0.77      0.78      0.77       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()
--------------------------------
{'clf__alpha': 0.5623413251903491, 'clf__fit_prior': True}
              precision    recall  f1-score   support

           0       0.85      0.78      0.81       103
           1       0.39      0.48      0.43        92
           2       0.42      0.32      0.36       101
           3       0.56      0.63      0.60       104

    accuracy                           0.56       400
   macro avg       0.56      0.55      0.55       400
weighted avg       0.56      0.56      0.55       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

1.5 PCA

1.5.1 On Raw data

[14]:
```
pca = PCA()
principalComp = pca.fit_transform(df)
plt.figure()
plt.plot(np.cumsum(pca.explained_variance_ratio_))
plt.xlabel('Number of components')
plt.ylabel('Variance(%)')
plt.title('Explained Variance')
plt.show()
n_components=4
pca = PCA(n_components=n_components)
new_data = pca.fit_transform(df)
df7 = pd.DataFrame(new_data,columns = [f'column {i}' for i in␣
 ↪range(n_components)])
df7
X_train, X_test, y_train, y_test =␣
 ↪train_test_split(df7,df2,random_state=5,shuffle=True,test_size=0.2)
Classification(X_train, X_test, y_train, y_test,data,'PCA_raw')
print('WITH GridSearch'+'\n'+"-"*80+'\n'+"-"*80 )
Classification(X_train, X_test, y_train,␣
 ↪y_test,data_GS,'PCA_raw_GS',Gridsearch=True)
```



```
-------------------------------
1. KNeighborsClassifier()
-------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.98      0.97       103
           1       0.88      0.92      0.90        92
           2       0.89      0.87      0.88       101
           3       0.97      0.92      0.95       104

    accuracy                           0.93       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.93      0.93      0.93       400

------------------------------------------------------------------------------
------------------------------------------------------------------------------
-------------------------------
2. LogisticRegression()
-------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.99      0.98       103
           1       0.94      0.95      0.94        92
           2       0.95      0.93      0.94       101
           3       0.98      0.96      0.97       104

    accuracy                           0.96       400
   macro avg       0.96      0.96      0.96       400
weighted avg       0.96      0.96      0.96       400

------------------------------------------------------------------------------
------------------------------------------------------------------------------
-------------------------------
3. DecisionTreeClassifier()
-------------------------------
{}
              precision    recall  f1-score   support

           0       0.92      0.93      0.93       103
           1       0.84      0.86      0.85        92
           2       0.85      0.85      0.85       101
           3       0.93      0.90      0.92       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
4. SVC()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.98      0.98      0.98       103
           1       0.89      0.96      0.92        92
           2       0.91      0.90      0.91       101
           3       0.99      0.93      0.96       104

    accuracy                           0.94       400
   macro avg       0.94      0.94      0.94       400
weighted avg       0.94      0.94      0.94       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
5. NuSVC()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.98      0.96      0.97       103
           1       0.85      0.93      0.89        92
           2       0.86      0.88      0.87       101
           3       0.99      0.89      0.94       104

    accuracy                           0.92       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
6. RandomForestClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.95      0.95       103
           1       0.88      0.93      0.91        92
           2       0.94      0.89      0.91       101
           3       0.96      0.95      0.96       104
    accuracy                           0.93       400
   macro avg       0.93      0.93      0.93       400
weighted avg       0.93      0.93      0.93       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.82      0.70      0.75       103
           1       0.56      0.72      0.63        92
           2       0.70      0.72      0.71       101
           3       0.92      0.79      0.85       104

    accuracy                           0.73       400
   macro avg       0.75      0.73      0.74       400
weighted avg       0.75      0.73      0.74       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.93      0.95       103
           1       0.84      0.95      0.89        92
           2       0.90      0.86      0.88       101
           3       0.95      0.91      0.93       104

    accuracy                           0.91       400
   macro avg       0.91      0.91      0.91       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.99      1.00      1.00       103
           1       0.51      0.72      0.59        92
           2       0.55      0.30      0.38       101
           3       0.92      0.98      0.95       104

    accuracy                           0.75       400
   macro avg       0.74      0.75      0.73       400
weighted avg       0.75      0.75      0.74       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.94      0.95       103
           1       0.70      0.71      0.70        92
           2       0.67      0.64      0.66       101
           3       0.87      0.90      0.89       104

    accuracy                           0.80       400
   macro avg       0.80      0.80      0.80       400
weighted avg       0.80      0.80      0.80       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.84      0.69      0.76       103
           1       0.36      0.49      0.41        92
           2       0.31      0.23      0.26       101
           3       0.52      0.58      0.55       104

    accuracy                           0.50       400
   macro avg       0.51      0.50      0.49       400
weighted avg       0.51      0.50      0.50       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
WITH GridSearch
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
1. KNeighborsClassifier()
```

```
--------------------------------
{'clf__algorithm': 'auto', 'clf__metric': 'euclidean', 'clf__n_neighbors': 22,
'clf__weights': 'distance'}
              precision    recall  f1-score   support

           0       0.98      0.98      0.98       103
           1       0.90      0.95      0.92        92
           2       0.92      0.91      0.92       101
           3       0.99      0.95      0.97       104

    accuracy                           0.95       400
   macro avg       0.95      0.95      0.95       400
weighted avg       0.95      0.95      0.95       400

----------------------------------------------------------------------------------
----------------------------------------------------------------------------------
--------------------------------
2. LogisticRegression()
--------------------------------
{'clf__C': 0.01, 'clf__penalty': 'l2', 'clf__solver': 'newton-cg'}
              precision    recall  f1-score   support

           0       1.00      0.98      0.99       103
           1       0.92      0.99      0.95        92
           2       0.95      0.91      0.93       101
           3       0.97      0.96      0.97       104

    accuracy                           0.96       400
   macro avg       0.96      0.96      0.96       400
weighted avg       0.96      0.96      0.96       400

----------------------------------------------------------------------------------
----------------------------------------------------------------------------------
--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{'clf__ccp_alpha': 0.0014384498882876629, 'clf__criterion': 'entropy',
'clf__max_features': None, 'clf__max_leaf_nodes': None,
'clf__min_samples_split': 2, 'clf__splitter': 'best'}
              precision    recall  f1-score   support

           0       0.94      0.94      0.94       103
           1       0.82      0.87      0.84        92
           2       0.85      0.81      0.83       101
           3       0.93      0.91      0.92       104

    accuracy                           0.89       400
   macro avg       0.88      0.88      0.88       400
```

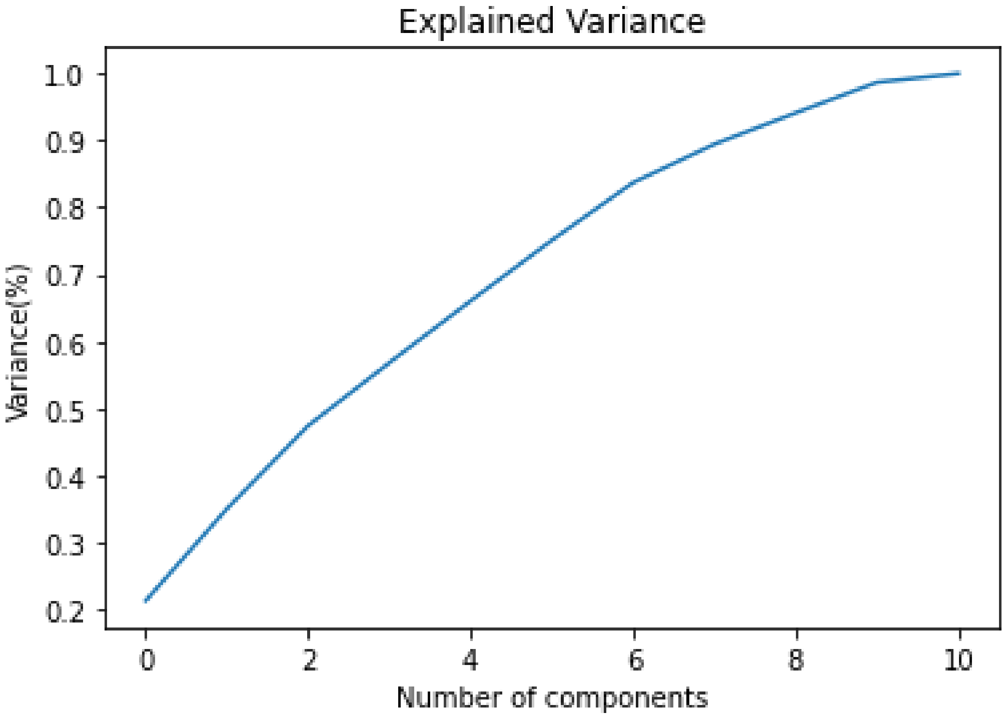
```
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()
--------------------------------
{'clf__C': 30, 'clf__decision_function_shape': 'ovo', 'clf__degree': 3,
'clf__kernel': 'linear'}
              precision    recall  f1-score   support

           0       1.00      0.98      0.99       103
           1       0.93      0.99      0.96        92
           2       0.94      0.93      0.94       101
           3       0.98      0.95      0.97       104

    accuracy                           0.96       400
   macro avg       0.96      0.96      0.96       400
weighted avg       0.96      0.96      0.96       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
5. NuSVC()
--------------------------------
{'clf__decision_function_shape': 'ovo', 'clf__degree': 3, 'clf__gamma': 'scale',
'clf__kernel': 'linear', 'clf__nu': 0.0707945784384138}
              precision    recall  f1-score   support

           0       1.00      0.98      0.99       103
           1       0.94      0.99      0.96        92
           2       0.95      0.94      0.95       101
           3       0.98      0.96      0.97       104

    accuracy                           0.97       400
   macro avg       0.97      0.97      0.97       400
weighted avg       0.97      0.97      0.97       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
6. RandomForestClassifier()
--------------------------------
{'clf__bootstrap': True, 'clf__criterion': 'gini', 'clf__max_depth': None,
'clf__max_features': 'log2', 'clf__n_estimators': 150}
              precision    recall  f1-score   support

           0       0.96      0.95      0.96       103
           1       0.89      0.92      0.90        92
           2       0.91      0.92      0.92       101
           3       0.98      0.94      0.96       104

    accuracy                           0.94       400
   macro avg       0.93      0.93      0.93       400
weighted avg       0.94      0.94      0.94       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
7. AdaBoostClassifier()
--------------------------------
{'clf__algorithm': 'SAMME', 'clf__base_estimator': RandomForestClassifier(),
'clf__n_estimators': 50}
              precision    recall  f1-score   support

           0       0.96      0.94      0.95       103
           1       0.88      0.93      0.91        92
           2       0.90      0.91      0.91       101
           3       0.97      0.92      0.95       104

    accuracy                           0.93       400
   macro avg       0.93      0.93      0.93       400
weighted avg       0.93      0.93      0.93       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{'clf__criterion': 'squared_error', 'clf__learning_rate': 0.31622776601683794,
'clf__loss': 'deviance', 'clf__max_features': 'log2', 'clf__n_estimators': 500}
              precision    recall  f1-score   support

           0       0.97      0.95      0.96       103
           1       0.86      0.95      0.90        92
           2       0.92      0.86      0.89       101
           3       0.95      0.94      0.95       104

    accuracy                           0.93       400
   macro avg       0.92      0.93      0.92       400
weighted avg       0.93      0.93      0.93       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
9. SGDClassifier()
```

```
--------------------------------
{'clf__alpha': 0.0032903445623126675, 'clf__learning_rate': 'optimal',
'clf__loss': 'squared_hinge', 'clf__penalty': 'l1'}
              precision    recall  f1-score   support

           0       1.00      0.98      0.99       103
           1       0.77      0.55      0.65        92
           2       0.66      0.84      0.74       101
           3       0.97      0.97      0.97       104

    accuracy                           0.84       400
   macro avg       0.85      0.84      0.84       400
weighted avg       0.85      0.84      0.84       400

---------------------------------------------------------------------------------
---------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()

--------------------------------
{'clf__var_smoothing': 1e-20}
              precision    recall  f1-score   support

           0       0.95      0.94      0.95       103
           1       0.70      0.71      0.70        92
           2       0.67      0.64      0.66       101
           3       0.87      0.90      0.89       104

    accuracy                           0.80       400
   macro avg       0.80      0.80      0.80       400
weighted avg       0.80      0.80      0.80       400

---------------------------------------------------------------------------------
---------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()

--------------------------------
{'clf__alpha': 0, 'clf__fit_prior': False}
              precision    recall  f1-score   support

           0       0.82      0.78      0.80       103
           1       0.41      0.40      0.41        92
           2       0.29      0.22      0.25       101
           3       0.49      0.63      0.55       104

    accuracy                           0.51       400
   macro avg       0.50      0.51      0.50       400
weighted avg       0.50      0.51      0.50       400
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

1.5.2 On df6(After VIF)

```
[15]: pca = PCA()
      principalComp = pca.fit_transform(df6)
      plt.figure()
      plt.plot(np.cumsum(pca.explained_variance_ratio_))
      plt.xlabel('Number of components')
      plt.ylabel('Variance(%)')
      plt.title('Explained Variance')
      plt.show()
      n_components=10
      pca = PCA(n_components=n_components)
      new_data = pca.fit_transform(df6)
      df7 = pd.DataFrame(new_data,columns = [f'column {i}' for i in␣
       ↪range(n_components)])
      df7
      X_train_ecvp, X_test_ecvp, y_train_ecvp, y_test_ecvp =␣
       ↪train_test_split(df7,df2,random_state=5,shuffle=True,test_size=0.2)
      Classification(X_train_ecvp, X_test_ecvp, y_train_ecvp, y_test_ecvp.values.
       ↪ravel(),data,'PCA_e+c+v')
      print('WITH GridSearch'+'\n'+"-"*80+'\n'+"-"*80 )
      Classification(X_train_ecvp, X_test_ecvp, y_train_ecvp, y_test_ecvp.values.
       ↪ravel(),data_GS,'PCA_e+c+v_GS',Gridsearch=True)
```



```
--------------------------------
1. KNeighborsClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.86      0.81      0.83       103
           1       0.56      0.70      0.62        92
           2       0.60      0.58      0.59       101
           3       0.85      0.74      0.79       104

    accuracy                           0.71       400
   macro avg       0.72      0.71      0.71       400
weighted avg       0.72      0.71      0.71       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
2. LogisticRegression()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.93      0.96      0.94       103
           1       0.83      0.82      0.82        92
           2       0.83      0.83      0.83       101
           3       0.94      0.92      0.93       104

    accuracy                           0.89       400
   macro avg       0.88      0.88      0.88       400
weighted avg       0.88      0.89      0.88       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
3. DecisionTreeClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.89      0.86      0.88       103
           1       0.67      0.77      0.72        92
           2       0.70      0.55      0.62       101
           3       0.79      0.87      0.83       104
    accuracy                           0.77       400
   macro avg       0.76      0.76      0.76       400
weighted avg       0.77      0.77      0.76       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.93      0.91      0.92       103
           1       0.79      0.79      0.79        92
           2       0.79      0.84      0.81       101
           3       0.94      0.89      0.92       104

    accuracy                           0.86       400
   macro avg       0.86      0.86      0.86       400
weighted avg       0.87      0.86      0.86       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
5. NuSVC()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.94      0.95       103
           1       0.84      0.82      0.83        92
           2       0.78      0.81      0.80       101
           3       0.89      0.89      0.89       104

    accuracy                           0.87       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.87      0.87      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
6. RandomForestClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.91      0.93      0.92       103
           1       0.73      0.78      0.76        92
           2       0.78      0.71      0.75       101
           3       0.90      0.91      0.91       104

    accuracy                           0.84       400
   macro avg       0.83      0.84      0.83       400
weighted avg       0.84      0.84      0.84       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.97      0.72      0.83       103
           1       0.52      0.82      0.64        92
           2       0.58      0.73      0.65       101
           3       0.98      0.51      0.67       104

    accuracy                           0.69       400
   macro avg       0.77      0.69      0.70       400
weighted avg       0.77      0.69      0.70       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.93      0.89      0.91       103
           1       0.70      0.78      0.74        92
           2       0.74      0.73      0.74       101
           3       0.93      0.88      0.90       104

    accuracy                           0.82       400
   macro avg       0.82      0.82      0.82       400
weighted avg       0.83      0.82      0.82       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()

--------------------------------
{}
              precision    recall  f1-score   support
           0       0.92      0.97      0.94       103
           1       0.56      0.38      0.45        92
           2       0.55      0.65      0.60       101
           3       0.92      0.96      0.94       104

    accuracy                           0.75       400
   macro avg       0.74      0.74      0.73       400
weighted avg       0.74      0.75      0.74       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.90      0.92      0.91       103
           1       0.66      0.65      0.66        92
           2       0.57      0.61      0.59       101
           3       0.79      0.71      0.75       104

    accuracy                           0.73       400
   macro avg       0.73      0.72      0.73       400
weighted avg       0.73      0.73      0.73       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.94      0.59      0.73       103
           1       0.44      0.84      0.58        92
           2       0.53      0.43      0.47       101
           3       0.84      0.64      0.73       104

    accuracy                           0.62       400
   macro avg       0.69      0.62      0.63       400
weighted avg       0.70      0.62      0.63       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
WITH GridSearch
--------------------------------------------------------------------------------
```

```
--------------------------------------------------------------------------------
--------------------------------
1. KNeighborsClassifier()
--------------------------------
{'clf__algorithm': 'auto', 'clf__metric': 'cityblock', 'clf__n_neighbors': 14,
'clf__weights': 'distance'}
              precision    recall  f1-score   support

           0       0.88      0.88      0.88       103
           1       0.67      0.72      0.69        92
           2       0.67      0.63      0.65       101
           3       0.82      0.81      0.82       104

    accuracy                           0.76       400
   macro avg       0.76      0.76      0.76       400
weighted avg       0.76      0.76      0.76       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
2. LogisticRegression()
--------------------------------
{'clf__C': 0.0117210229753348, 'clf__penalty': 'none', 'clf__solver': 'saga'}
              precision    recall  f1-score   support

           0       0.93      0.96      0.94       103
           1       0.83      0.82      0.82        92
           2       0.84      0.83      0.84       101
           3       0.94      0.93      0.94       104

    accuracy                           0.89       400
   macro avg       0.89      0.89      0.89       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{'clf__ccp_alpha': 0.001623776739188721, 'clf__criterion': 'gini',
'clf__max_features': None, 'clf__max_leaf_nodes': None,
'clf__min_samples_split': 2, 'clf__splitter': 'best'}
              precision    recall  f1-score   support

           0       0.87      0.87      0.87       103
           1       0.64      0.70      0.67        92
           2       0.69      0.58      0.63       101
           3       0.81      0.88      0.84       104
    accuracy                           0.76       400
   macro avg       0.76      0.76      0.75       400
weighted avg       0.76      0.76      0.76       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()
--------------------------------
{'clf__C': 0.8, 'clf__decision_function_shape': 'ovo', 'clf__degree': 3,
'clf__kernel': 'linear'}
              precision    recall  f1-score   support

           0       0.92      0.95      0.93       103
           1       0.81      0.80      0.81        92
           2       0.82      0.82      0.82       101
           3       0.94      0.91      0.93       104

    accuracy                           0.88       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.87      0.88      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
5. NuSVC()
--------------------------------
{'clf__decision_function_shape': 'ovo', 'clf__degree': 3, 'clf__gamma': 'scale',
'clf__kernel': 'linear', 'clf__nu': 0.0707945784384138}
              precision    recall  f1-score   support

           0       0.94      0.88      0.91       103
           1       0.73      0.85      0.78        92
           2       0.77      0.73      0.75       101
           3       0.90      0.87      0.88       104

    accuracy                           0.83       400
   macro avg       0.83      0.83      0.83       400
weighted avg       0.84      0.83      0.83       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
6. RandomForestClassifier()
--------------------------------
{'clf__bootstrap': True, 'clf__criterion': 'entropy', 'clf__max_depth': None,
'clf__max_features': 'sqrt', 'clf__n_estimators': 200}
```

```
              precision    recall  f1-score   support

           0       0.92      0.93      0.93       103
           1       0.76      0.78      0.77        92
           2       0.78      0.74      0.76       101
           3       0.90      0.91      0.91       104

    accuracy                           0.84       400
   macro avg       0.84      0.84      0.84       400
weighted avg       0.84      0.84      0.84       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
7. AdaBoostClassifier()
--------------------------------
{'clf__algorithm': 'SAMME.R', 'clf__base_estimator': RandomForestClassifier(),
'clf__n_estimators': 100}
              precision    recall  f1-score   support

           0       0.92      0.95      0.93       103
           1       0.78      0.80      0.79        92
           2       0.79      0.71      0.75       101
           3       0.88      0.90      0.89       104

    accuracy                           0.84       400
   macro avg       0.84      0.84      0.84       400
weighted avg       0.84      0.84      0.84       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{'clf__criterion': 'friedman_mse', 'clf__learning_rate': 0.1, 'clf__loss':
'deviance', 'clf__max_features': 'log2', 'clf__n_estimators': 900}
              precision    recall  f1-score   support

           0       0.91      0.91      0.91       103
           1       0.70      0.77      0.74        92
           2       0.75      0.71      0.73       101
           3       0.92      0.88      0.90       104

    accuracy                           0.82       400
   macro avg       0.82      0.82      0.82       400
weighted avg       0.83      0.82      0.82       400

--------------------------------------------------------------------------------
```

```
--------------------------------------------------------------------------------
--------------------------------
9. SGDClassifier()
--------------------------------
{'clf__alpha': 0.0009236708571873865, 'clf__learning_rate': 'optimal',
'clf__loss': 'log', 'clf__penalty': 'l1'}
              precision    recall  f1-score   support

           0       0.93      0.96      0.95       103
           1       0.63      0.62      0.63        92
           2       0.66      0.66      0.66       101
           3       0.94      0.93      0.94       104

    accuracy                           0.80       400
   macro avg       0.79      0.79      0.79       400
weighted avg       0.80      0.80      0.80       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
10. GaussianNB()
--------------------------------
{'clf__var_smoothing': 1e-20}
              precision    recall  f1-score   support

           0       0.90      0.92      0.91       103
           1       0.66      0.65      0.66        92
           2       0.57      0.61      0.59       101
           3       0.79      0.71      0.75       104

    accuracy                           0.73       400
   macro avg       0.73      0.72      0.73       400
weighted avg       0.73      0.73      0.73       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
11. MultinomialNB()
--------------------------------
{'clf__alpha': 0, 'clf__fit_prior': False}
              precision    recall  f1-score   support

           0       0.86      0.76      0.80       103
           1       0.51      0.59      0.55        92
           2       0.54      0.56      0.55       101
           3       0.77      0.73      0.75       104

    accuracy                           0.66       400
   macro avg       0.67      0.66      0.66       400
weighted avg       0.68      0.66      0.67       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

1.6 Reverse Feature Elimination

```
[17]: X_train_e, X_test_e, y_train_e, y_test_e =␣
       ↪train_test_split(df4,df2,random_state=5,shuffle=True,test_size=0.2)
```

```
[18]: rfe = RFE(GradientBoostingClassifier(random_state=42),
                n_features_to_select=10)
      rfe.fit(X_train_e, y_train_e.values.ravel())
      rfe_features = X_train_e.columns[rfe.support_]
      dfr = df4[rfe_features]
      print(f"===== {len(rfe_features)} features were selected =====")
      print(f"{', '.join(rfe_features)}")
      X_train_er, X_test_er, y_train_er, y_test_er =␣
       ↪train_test_split(dfr,df2,random_state=5,shuffle=True,test_size=0.2)

      Classification(X_train_er, X_test_er, y_train_er, y_test_er,data,'RFE_e')
      print('WITH GridSearch'+'\n'+"-"*80+'\n'+"-"*80 )
      Classification(X_train_er, X_test_er, y_train_er,␣
       ↪y_test_er,data_GS,'RFE_e_GS',Gridsearch=True)
```

```
===== 10 features were selected =====
battery_power, mobile_wt, px_height, px_width, ram, px_diag, px_area,
talk_per_mAh, ram_per_core, weight_by_thickness

--------------------------------
1. KNeighborsClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.43      0.50      0.47       103
           1       0.24      0.33      0.28        92
           2       0.30      0.28      0.29       101
           3       0.30      0.17      0.22       104

    accuracy                           0.32       400
   macro avg       0.32      0.32      0.31       400
weighted avg       0.32      0.32      0.31       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
2. LogisticRegression()
```

```
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.89      0.85      0.87       103
           1       0.46      0.49      0.48        92
           2       0.29      0.24      0.26       101
           3       0.60      0.69      0.64       104

    accuracy                           0.57       400
   macro avg       0.56      0.57      0.56       400
weighted avg       0.56      0.57      0.57       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
3. DecisionTreeClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.94      0.92      0.93       103
           1       0.84      0.87      0.86        92
           2       0.83      0.84      0.84       101
           3       0.91      0.89      0.90       104

    accuracy                           0.88       400
   macro avg       0.88      0.88      0.88       400
weighted avg       0.88      0.88      0.88       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.98      0.96      0.97       103
           1       0.85      0.92      0.89        92
           2       0.88      0.84      0.86       101
           3       0.95      0.93      0.94       104

    accuracy                           0.92       400
   macro avg       0.91      0.91      0.91       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
```

```
--------------------------------------------------------------------------------
--------------------------------
5. NuSVC()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.95      0.96       103
           1       0.84      0.89      0.86        92
           2       0.87      0.84      0.85       101
           3       0.95      0.93      0.94       104

    accuracy                           0.91       400
   macro avg       0.90      0.90      0.90       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
6. RandomForestClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.99      0.96      0.98       103
           1       0.86      0.90      0.88        92
           2       0.84      0.89      0.87       101
           3       0.98      0.91      0.95       104

    accuracy                           0.92       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
7. AdaBoostClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.92      0.55      0.69       103
           1       0.58      0.80      0.67        92
           2       0.63      0.85      0.73       101
           3       0.91      0.64      0.75       104

    accuracy                           0.71       400
   macro avg       0.76      0.71      0.71       400
```

```
weighted avg       0.76      0.71      0.71       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.97      0.96      0.97       103
           1       0.84      0.93      0.89        92
           2       0.90      0.85      0.87       101
           3       0.97      0.93      0.95       104

    accuracy                           0.92       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.99      0.97       103
           1       0.54      0.48      0.51        92
           2       0.55      0.59      0.57       101
           3       0.97      0.94      0.96       104

    accuracy                           0.76       400
   macro avg       0.75      0.75      0.75       400
weighted avg       0.76      0.76      0.76       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.92      0.95      0.94       103
           1       0.68      0.67      0.68        92
           2       0.66      0.70      0.68       101
           3       0.93      0.85      0.88       104
    accuracy                           0.80       400
   macro avg       0.80      0.79      0.79       400
weighted avg       0.80      0.80      0.80       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
11. MultinomialNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.85      0.70      0.77       103
           1       0.40      0.46      0.42        92
           2       0.45      0.50      0.47       101
           3       0.68      0.64      0.66       104

    accuracy                           0.58       400
   macro avg       0.59      0.57      0.58       400
weighted avg       0.60      0.58      0.59       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
WITH GridSearch
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
1. KNeighborsClassifier()
--------------------------------
{'clf__algorithm': 'auto', 'clf__metric': 'cosine', 'clf__n_neighbors': 9,
'clf__weights': 'distance'}
              precision    recall  f1-score   support

           0       0.92      0.83      0.87       103
           1       0.55      0.51      0.53        92
           2       0.44      0.49      0.46       101
           3       0.63      0.68      0.66       104

    accuracy                           0.63       400
   macro avg       0.64      0.63      0.63       400
weighted avg       0.64      0.63      0.63       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
2. LogisticRegression()
--------------------------------
```

```
{'clf__C': 0.024920211513780568, 'clf__penalty': 'l2', 'clf__solver': 'newton-
cg'}
              precision    recall  f1-score   support

           0       0.91      0.92      0.92       103
           1       0.71      0.65      0.68        92
           2       0.65      0.67      0.66       101
           3       0.86      0.88      0.87       104

    accuracy                           0.79       400
   macro avg       0.78      0.78      0.78       400
weighted avg       0.78      0.79      0.78       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{'clf__ccp_alpha': 0.003359818286283781, 'clf__criterion': 'entropy',
'clf__max_features': None, 'clf__max_leaf_nodes': None,
'clf__min_samples_split': 12, 'clf__splitter': 'best'}
              precision    recall  f1-score   support

           0       0.94      0.94      0.94       103
           1       0.79      0.91      0.85        92
           2       0.84      0.76      0.80       101
           3       0.92      0.88      0.90       104

    accuracy                           0.87       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.88      0.87      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
4. SVC()
--------------------------------
{'clf__C': 30, 'clf__decision_function_shape': 'ovo', 'clf__degree': 3,
'clf__kernel': 'linear'}
              precision    recall  f1-score   support

           0       1.00      0.99      1.00       103
           1       0.94      0.99      0.96        92
           2       0.96      0.94      0.95       101
           3       0.99      0.97      0.98       104

    accuracy                           0.97       400
   macro avg       0.97      0.97      0.97       400
```

```
weighted avg       0.97      0.97      0.97       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
5. NuSVC()
--------------------------------
{'clf__decision_function_shape': 'ovo', 'clf__degree': 3, 'clf__gamma': 'scale',
'clf__kernel': 'linear', 'clf__nu': 0.06760829753919818}
              precision    recall  f1-score   support

           0       1.00      0.99      1.00       103
           1       0.93      0.99      0.96        92
           2       0.96      0.93      0.94       101
           3       0.99      0.97      0.98       104

    accuracy                           0.97       400
   macro avg       0.97      0.97      0.97       400
weighted avg       0.97      0.97      0.97       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
6. RandomForestClassifier()
--------------------------------
{'clf__bootstrap': True, 'clf__criterion': 'entropy', 'clf__max_depth': None,
'clf__max_features': None, 'clf__n_estimators': 200}
              precision    recall  f1-score   support

           0       0.96      0.95      0.96       103
           1       0.89      0.95      0.92        92
           2       0.91      0.92      0.92       101
           3       0.98      0.92      0.95       104

    accuracy                           0.94       400
   macro avg       0.93      0.94      0.93       400
weighted avg       0.94      0.94      0.94       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()
--------------------------------
{'clf__algorithm': 'SAMME.R', 'clf__base_estimator': RandomForestClassifier(),
'clf__n_estimators': 100}
              precision    recall  f1-score   support

           0       1.00      0.95      0.98       103
           1       0.85      0.96      0.90        92
           2       0.88      0.88      0.88       101
           3       0.98      0.92      0.95       104

    accuracy                           0.93       400
   macro avg       0.93      0.93      0.93       400
weighted avg       0.93      0.93      0.93       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{'clf__criterion': 'squared_error', 'clf__learning_rate': 0.31622776601683794,
'clf__loss': 'deviance', 'clf__max_features': 'log2', 'clf__n_estimators': 300}
              precision    recall  f1-score   support

           0       0.99      0.97      0.98       103
           1       0.86      0.93      0.90        92
           2       0.85      0.84      0.85       101
           3       0.95      0.90      0.93       104

    accuracy                           0.91       400
   macro avg       0.91      0.91      0.91       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()
--------------------------------
{'clf__alpha': 0.0012689610031679222, 'clf__learning_rate': 'optimal',
'clf__loss': 'log', 'clf__penalty': 'l1'}
              precision    recall  f1-score   support

           0       1.00      0.98      0.99       103
           1       0.69      0.65      0.67        92
           2       0.68      0.72      0.70       101
           3       0.97      0.98      0.98       104

    accuracy                           0.84       400
   macro avg       0.84      0.83      0.83       400
weighted avg       0.84      0.84      0.84       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()
```

```
--------------------------------
{'clf__var_smoothing': 1e-20}
              precision    recall  f1-score   support

           0       0.92      0.95      0.94       103
           1       0.68      0.67      0.68        92
           2       0.66      0.70      0.68       101
           3       0.93      0.85      0.88       104

    accuracy                           0.80       400
   macro avg       0.80      0.79      0.79       400
weighted avg       0.80      0.80      0.80       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()

--------------------------------
{'clf__alpha': 31.622776601683793, 'clf__fit_prior': True}
              precision    recall  f1-score   support

           0       0.82      0.76      0.79       103
           1       0.44      0.46      0.45        92
           2       0.48      0.46      0.47       101
           3       0.67      0.73      0.70       104

    accuracy                           0.60       400
   macro avg       0.60      0.60      0.60       400
weighted avg       0.61      0.60      0.61       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

1.7 Reverse Feature Elimination with CV

[19]:
```
rfecv = RFECV(GradientBoostingClassifier(random_state=42),
          min_features_to_select=10,cv=10,n_jobs=-1)
rfecv.fit(X_train_e, y_train_e.values.ravel())
rfecv_features = X_train_e.columns[rfe.support_]
dfrv = df4[rfecv_features]
print(f"===== {len(rfecv_features)} features were selected =====")
print(f"{', '.join(rfecv_features)}")
X_train_er, X_test_er, y_train_er, y_test_er =␣
 ↪train_test_split(dfrv,df2,random_state=5,shuffle=True,test_size=0.2)

Classification(X_train_er, X_test_er, y_train_er, y_test_er,data,'RFECV_e')
print('WITH GridSearch'+'\n'+"-"*80+'\n'+"-"*80 )
```

```
Classification(X_train_er, X_test_er, y_train_er,␣
 ↪y_test_er,data_GS,'RFECV_e_GS',Gridsearch=True)
```

```
===== 10 features were selected =====
battery_power, mobile_wt, px_height, px_width, ram, px_diag, px_area,
talk_per_mAh, ram_per_core, weight_by_thickness
--------------------------------
1. KNeighborsClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.43      0.50      0.47       103
           1       0.24      0.33      0.28        92
           2       0.30      0.28      0.29       101
           3       0.30      0.17      0.22       104

    accuracy                           0.32       400
   macro avg       0.32      0.32      0.31       400
weighted avg       0.32      0.32      0.31       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
2. LogisticRegression()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.89      0.85      0.87       103
           1       0.46      0.49      0.48        92
           2       0.29      0.24      0.26       101
           3       0.60      0.69      0.64       104

    accuracy                           0.57       400
   macro avg       0.56      0.57      0.56       400
weighted avg       0.56      0.57      0.57       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.92      0.94       103
           1       0.84      0.88      0.86        92
           2       0.83      0.84      0.84       101
           3       0.91      0.89      0.90       104

    accuracy                           0.89       400
   macro avg       0.88      0.88      0.88       400
weighted avg       0.89      0.89      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.98      0.96      0.97       103
           1       0.85      0.92      0.89        92
           2       0.88      0.84      0.86       101
           3       0.95      0.93      0.94       104

    accuracy                           0.92       400
   macro avg       0.91      0.91      0.91       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
5. NuSVC()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.96      0.95      0.96       103
           1       0.84      0.89      0.86        92
           2       0.87      0.84      0.85       101
           3       0.95      0.93      0.94       104

    accuracy                           0.91       400
   macro avg       0.90      0.90      0.90       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
6. RandomForestClassifier()

--------------------------------
{}
              precision    recall  f1-score   support
           0       0.98      0.96      0.97       103
           1       0.87      0.92      0.89        92
           2       0.87      0.89      0.88       101
           3       0.98      0.92      0.95       104

    accuracy                           0.93       400
   macro avg       0.93      0.92      0.92       400
weighted avg       0.93      0.93      0.93       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
7. AdaBoostClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.92      0.55      0.69       103
           1       0.58      0.80      0.67        92
           2       0.63      0.85      0.73       101
           3       0.91      0.64      0.75       104

    accuracy                           0.71       400
   macro avg       0.76      0.71      0.71       400
weighted avg       0.76      0.71      0.71       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.97      0.96      0.97       103
           1       0.84      0.93      0.89        92
           2       0.90      0.85      0.87       101
           3       0.97      0.93      0.95       104

    accuracy                           0.92       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
9. SGDClassifier()
```

```
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.94      0.98      0.96       103
           1       0.45      0.53      0.49        92
           2       0.52      0.45      0.48       101
           3       0.98      0.91      0.95       104

    accuracy                           0.73       400
   macro avg       0.72      0.72      0.72       400
weighted avg       0.73      0.72      0.73       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
10. GaussianNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.92      0.95      0.94       103
           1       0.68      0.67      0.68        92
           2       0.66      0.70      0.68       101
           3       0.93      0.85      0.88       104

    accuracy                           0.80       400
   macro avg       0.80      0.79      0.79       400
weighted avg       0.80      0.80      0.80       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
11. MultinomialNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.85      0.70      0.77       103
           1       0.40      0.46      0.42        92
           2       0.45      0.50      0.47       101
           3       0.68      0.64      0.66       104

    accuracy                           0.58       400
   macro avg       0.59      0.57      0.58       400
weighted avg       0.60      0.58      0.59       400

--------------------------------------------------------------------------------
```

```
--------------------------------------------------------------------------------
WITH GridSearch
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
________________________________
1. KNeighborsClassifier()
________________________________
{'clf__algorithm': 'auto', 'clf__metric': 'cosine', 'clf__n_neighbors': 9,
'clf__weights': 'distance'}
              precision    recall  f1-score   support

           0       0.92      0.83      0.87       103
           1       0.55      0.51      0.53        92
           2       0.44      0.49      0.46       101
           3       0.63      0.68      0.66       104

    accuracy                           0.63       400
   macro avg       0.64      0.63      0.63       400
weighted avg       0.64      0.63      0.63       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
________________________________
2. LogisticRegression()
________________________________
{'clf__C': 0.02492021151378O568, 'clf__penalty': 'l2', 'clf__solver': 'newton-
cg'}
              precision    recall  f1-score   support

           0       0.91      0.92      0.92       103
           1       0.71      0.65      0.68        92
           2       0.65      0.67      0.66       101
           3       0.86      0.88      0.87       104

    accuracy                           0.79       400
   macro avg       0.78      0.78      0.78       400
weighted avg       0.78      0.79      0.78       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
________________________________
3. DecisionTreeClassifier()
________________________________
{'clf__ccp_alpha': 0.0012742749857031334, 'clf__criterion': 'entropy',
'clf__max_features': None, 'clf__max_leaf_nodes': None,
'clf__min_samples_split': 2, 'clf__splitter': 'best'}
              precision    recall  f1-score   support
           0       0.95      0.94      0.95       103
           1       0.83      0.93      0.88        92
           2       0.88      0.80      0.84       101
           3       0.92      0.90      0.91       104

    accuracy                           0.90       400
   macro avg       0.89      0.90      0.89       400
weighted avg       0.90      0.90      0.89       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
4. SVC()
--------------------------------
{'clf__C': 30, 'clf__decision_function_shape': 'ovo', 'clf__degree': 3,
'clf__kernel': 'linear'}
              precision    recall  f1-score   support

           0       1.00      0.99      1.00       103
           1       0.94      0.99      0.96        92
           2       0.96      0.94      0.95       101
           3       0.99      0.97      0.98       104

    accuracy                           0.97       400
   macro avg       0.97      0.97      0.97       400
weighted avg       0.97      0.97      0.97       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
5. NuSVC()
--------------------------------
{'clf__decision_function_shape': 'ovo', 'clf__degree': 3, 'clf__gamma': 'scale',
'clf__kernel': 'linear', 'clf__nu': 0.06760829753919818}
              precision    recall  f1-score   support

           0       1.00      0.99      1.00       103
           1       0.93      0.99      0.96        92
           2       0.96      0.93      0.94       101
           3       0.99      0.97      0.98       104

    accuracy                           0.97       400
   macro avg       0.97      0.97      0.97       400
weighted avg       0.97      0.97      0.97       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
```

```
6. RandomForestClassifier()
--------------------------------
{'clf__bootstrap': True, 'clf__criterion': 'entropy', 'clf__max_depth': None,
 'clf__max_features': 'sqrt', 'clf__n_estimators': 150}
              precision    recall  f1-score   support

           0       0.98      0.96      0.97       103
           1       0.87      0.91      0.89        92
           2       0.89      0.90      0.90       101
           3       0.99      0.95      0.97       104

    accuracy                           0.93       400
   macro avg       0.93      0.93      0.93       400
weighted avg       0.93      0.93      0.93       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()
--------------------------------
{'clf__algorithm': 'SAMME.R', 'clf__base_estimator': RandomForestClassifier(),
 'clf__n_estimators': 100}
              precision    recall  f1-score   support

           0       0.99      0.96      0.98       103
           1       0.85      0.90      0.87        92
           2       0.84      0.87      0.85       101
           3       0.98      0.91      0.95       104

    accuracy                           0.91       400
   macro avg       0.91      0.91      0.91       400
weighted avg       0.92      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{'clf__criterion': 'friedman_mse', 'clf__learning_rate': 1.0, 'clf__loss':
'deviance', 'clf__max_features': 'log2', 'clf__n_estimators': 700}
              precision    recall  f1-score   support

           0       0.98      0.95      0.97       103
           1       0.86      0.92      0.89        92
           2       0.87      0.87      0.87       101
           3       0.96      0.92      0.94       104

    accuracy                           0.92       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
9. SGDClassifier()
--------------------------------
{'clf__alpha': 0.002395026619987486, 'clf__learning_rate': 'optimal',
'clf__loss': 'log', 'clf__penalty': 'l1'}
              precision    recall  f1-score   support

           0       0.97      0.98      0.98       103
           1       0.65      0.74      0.69        92
           2       0.74      0.62      0.68       101
           3       0.97      0.99      0.98       104

    accuracy                           0.84       400
   macro avg       0.83      0.83      0.83       400
weighted avg       0.84      0.84      0.84       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
10. GaussianNB()
--------------------------------
{'clf__var_smoothing': 1e-20}
              precision    recall  f1-score   support

           0       0.92      0.95      0.94       103
           1       0.68      0.67      0.68        92
           2       0.66      0.70      0.68       101
           3       0.93      0.85      0.88       104

    accuracy                           0.80       400
   macro avg       0.80      0.79      0.79       400
weighted avg       0.80      0.80      0.80       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
11. MultinomialNB()
--------------------------------
{'clf__alpha': 31.622776601683793, 'clf__fit_prior': True}
              precision    recall  f1-score   support

           0       0.82      0.76      0.79       103
           1       0.44      0.46      0.45        92
           2       0.48      0.46      0.47       101
           3       0.67      0.73      0.70       104

    accuracy                           0.60       400
   macro avg       0.60      0.60      0.60       400
weighted avg       0.61      0.60      0.61       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

1.8 SelectKBest

[20]:

```
kbest = SelectKBest(k=10)
kbest.fit(X_train_e, y_train_e.values.ravel())
# See selected features
kbest_features = X_train_e.columns[kbest.get_support()]
dfk = df4[kbest_features]
print(f"===== {len(kbest_features)} features were selected =====")
print(f"{', '.join(kbest_features)}")
X_train_ek, X_test_ek, y_train_ek, y_test_ek =␣
 ↪train_test_split(dfk,df2,random_state=5,shuffle=True,test_size=0.2)

Classification(X_train_ek, X_test_ek, y_train_ek, y_test_ek,data,'KBest_e')
print('WITH GridSearch'+'\n'+"-"*80+'\n'+"-"*80 )
Classification(X_train_ek, X_test_ek, y_train_ek,␣
 ↪y_test_ek,data_GS,'KBest_e_GS',Gridsearch=True)
```

```
===== 10 features were selected =====
battery_power, px_height, px_width, ram, px_diag, px_area, talk_per_mAh,
ram_per_core, sc_res, DPIy

--------------------------------
1. KNeighborsClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.45      0.53      0.49       103
           1       0.23      0.32      0.27        92
           2       0.30      0.28      0.29       101
           3       0.30      0.17      0.22       104

    accuracy                           0.33       400
   macro avg       0.32      0.32      0.32       400
weighted avg       0.32      0.33      0.32       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
```

```
2. LogisticRegression()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.95      0.69      0.80       103
           1       0.43      0.57      0.49        92
           2       0.29      0.21      0.24       101
           3       0.58      0.73      0.65       104

    accuracy                           0.55       400
   macro avg       0.56      0.55      0.54       400
weighted avg       0.57      0.55      0.55       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.94      0.93      0.94       103
           1       0.89      0.86      0.87        92
           2       0.84      0.88      0.86       101
           3       0.91      0.90      0.91       104

    accuracy                           0.90       400
   macro avg       0.90      0.89      0.89       400
weighted avg       0.90      0.90      0.90       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
4. SVC()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.99      0.97      0.98       103
           1       0.86      0.93      0.90        92
           2       0.89      0.85      0.87       101
           3       0.96      0.94      0.95       104

    accuracy                           0.93       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.93      0.93      0.93       400
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
5. NuSVC()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.97      0.96      0.97       103
           1       0.86      0.90      0.88        92
           2       0.87      0.84      0.85       101
           3       0.94      0.93      0.94       104

    accuracy                           0.91       400
   macro avg       0.91      0.91      0.91       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
6. RandomForestClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.99      0.96      0.98       103
           1       0.87      0.89      0.88        92
           2       0.85      0.90      0.88       101
           3       0.98      0.93      0.96       104

    accuracy                           0.92       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.93      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.83      0.29      0.43       103
           1       0.46      0.83      0.59        92
           2       0.67      0.71      0.69       101
           3       0.86      0.75      0.80       104

    accuracy                           0.64       400
   macro avg       0.70      0.65      0.63       400
weighted avg       0.71      0.64      0.63       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
8. GradientBoostingClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.97      0.96      0.97       103
           1       0.83      0.92      0.88        92
           2       0.89      0.82      0.86       101
           3       0.95      0.94      0.95       104

    accuracy                           0.91       400
   macro avg       0.91      0.91      0.91       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
9. SGDClassifier()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.97      0.99      0.98       103
           1       0.60      0.85      0.70        92
           2       0.77      0.47      0.58       101
           3       0.97      0.97      0.97       104

    accuracy                           0.82       400
   macro avg       0.83      0.82      0.81       400
weighted avg       0.84      0.82      0.81       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
10. GaussianNB()

--------------------------------
{}
              precision    recall  f1-score   support

           0       0.92      0.93      0.93       103
           1       0.65      0.64      0.64        92
           2       0.61      0.67      0.64       101
           3       0.91      0.83      0.87       104

    accuracy                           0.77       400
   macro avg       0.77      0.77      0.77       400
weighted avg       0.78      0.77      0.77       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()
--------------------------------
{}
              precision    recall  f1-score   support

           0       0.82      0.71      0.76       103
           1       0.36      0.41      0.38        92
           2       0.42      0.47      0.44       101
           3       0.64      0.57      0.60       104

    accuracy                           0.54       400
   macro avg       0.56      0.54      0.55       400
weighted avg       0.57      0.54      0.55       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
WITH GridSearch
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
1. KNeighborsClassifier()
--------------------------------
{'clf__algorithm': 'auto', 'clf__metric': 'cosine', 'clf__n_neighbors': 9,
'clf__weights': 'distance'}
              precision    recall  f1-score   support

           0       0.89      0.87      0.88       103
           1       0.58      0.54      0.56        92
           2       0.47      0.50      0.49       101
           3       0.67      0.67      0.67       104

    accuracy                           0.65       400
   macro avg       0.65      0.65      0.65       400
weighted avg       0.66      0.65      0.65       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
2. LogisticRegression()
```

```
--------------------------------
{'clf__C': 0.02592943797404667, 'clf__penalty': 'l2', 'clf__solver': 'newton-
cg'}
              precision    recall  f1-score   support

           0       0.94      0.90      0.92       103
           1       0.71      0.65      0.68        92
           2       0.56      0.59      0.58       101
           3       0.80      0.84      0.82       104

    accuracy                           0.75       400
   macro avg       0.75      0.75      0.75       400
weighted avg       0.75      0.75      0.75       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
3. DecisionTreeClassifier()
--------------------------------
{'clf__ccp_alpha': 0.0011288378916846896, 'clf__criterion': 'entropy',
'clf__max_features': None, 'clf__max_leaf_nodes': None,
'clf__min_samples_split': 2, 'clf__splitter': 'random'}
              precision    recall  f1-score   support

           0       0.94      0.93      0.94       103
           1       0.78      0.82      0.80        92
           2       0.81      0.80      0.81       101
           3       0.94      0.92      0.93       104

    accuracy                           0.87       400
   macro avg       0.87      0.87      0.87       400
weighted avg       0.87      0.87      0.87       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
4. SVC()
--------------------------------
{'clf__C': 2, 'clf__decision_function_shape': 'ovo', 'clf__degree': 3,
'clf__kernel': 'linear'}
              precision    recall  f1-score   support

           0       1.00      0.97      0.99       103
           1       0.91      1.00      0.95        92
           2       0.97      0.92      0.94       101
           3       0.98      0.97      0.98       104

    accuracy                           0.96       400
   macro avg       0.97      0.97      0.96       400
weighted avg       0.97      0.96      0.97       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
5. NuSVC()
--------------------------------
{'clf__decision_function_shape': 'ovo', 'clf__degree': 3, 'clf__gamma': 'scale',
'clf__kernel': 'linear', 'clf__nu': 0.06760829753919818}
              precision    recall  f1-score   support

           0       1.00      0.98      0.99       103
           1       0.91      0.99      0.95        92
           2       0.95      0.92      0.93       101
           3       0.99      0.96      0.98       104

    accuracy                           0.96       400
   macro avg       0.96      0.96      0.96       400
weighted avg       0.96      0.96      0.96       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
6. RandomForestClassifier()
--------------------------------
{'clf__bootstrap': True, 'clf__criterion': 'entropy', 'clf__max_depth': None,
'clf__max_features': None, 'clf__n_estimators': 200}
              precision    recall  f1-score   support

           0       0.96      0.94      0.95       103
           1       0.86      0.95      0.90        92
           2       0.94      0.88      0.91       101
           3       0.96      0.95      0.96       104

    accuracy                           0.93       400
   macro avg       0.93      0.93      0.93       400
weighted avg       0.93      0.93      0.93       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
7. AdaBoostClassifier()
--------------------------------
{'clf__algorithm': 'SAMME.R', 'clf__base_estimator': RandomForestClassifier(),
'clf__n_estimators': 100}
              precision    recall  f1-score   support
           0       0.99      0.95      0.97       103
           1       0.86      0.89      0.88        92
           2       0.85      0.90      0.88       101
           3       0.98      0.93      0.96       104

    accuracy                           0.92       400
   macro avg       0.92      0.92      0.92       400
weighted avg       0.92      0.92      0.92       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
8. GradientBoostingClassifier()
--------------------------------
{'clf__criterion': 'squared_error', 'clf__learning_rate': 1.0, 'clf__loss':
'deviance', 'clf__max_features': 'log2', 'clf__n_estimators': 500}
              precision    recall  f1-score   support

           0       0.99      0.95      0.97       103
           1       0.84      0.92      0.88        92
           2       0.85      0.83      0.84       101
           3       0.94      0.91      0.93       104

    accuracy                           0.91       400
   macro avg       0.91      0.91      0.90       400
weighted avg       0.91      0.91      0.91       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
9. SGDClassifier()
--------------------------------
{'clf__alpha': 0.01, 'clf__learning_rate': 'optimal', 'clf__loss':
'modified_huber', 'clf__penalty': 'l1'}
              precision    recall  f1-score   support

           0       0.99      0.99      0.99       103
           1       0.72      0.82      0.77        92
           2       0.77      0.68      0.72       101
           3       0.96      0.95      0.96       104

    accuracy                           0.86       400
   macro avg       0.86      0.86      0.86       400
weighted avg       0.86      0.86      0.86       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------
```

```
10. GaussianNB()

--------------------------------
{'clf__var_smoothing': 1e-20}
              precision    recall  f1-score   support

           0       0.92      0.93      0.93       103
           1       0.65      0.64      0.64        92
           2       0.61      0.67      0.64       101
           3       0.91      0.83      0.87       104

    accuracy                           0.77       400
   macro avg       0.77      0.77      0.77       400
weighted avg       0.78      0.77      0.77       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------

--------------------------------
11. MultinomialNB()

--------------------------------
{'clf__alpha': 0.5623413251903491, 'clf__fit_prior': True}
              precision    recall  f1-score   support

           0       0.82      0.71      0.76       103
           1       0.36      0.41      0.38        92
           2       0.42      0.47      0.44       101
           3       0.64      0.57      0.60       104

    accuracy                           0.54       400
   macro avg       0.56      0.54      0.55       400
weighted avg       0.57      0.54      0.55       400

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
```

```
[21]: pd.DataFrame(data).to_csv('data.csv')
      pd.DataFrame(data_GS).to_csv('data_GS.csv')
```

```
[ ]:
```