Στον 0ο διαχωρισμό του dataset έχουμε:

```
################################################################################################################
########################
######                                  GAUSSIAN_NB MODEL (0ος διαχωρισμος)
######
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=1.873817422860383e-05) {'var_smoothing': 1.873817422860383e-05}
The accuracy in training set is: 84.87%
The precision in training set is: 89.74%
The recall in training set is: 90.52%
The F1 score in training set is: 90.13%
The confusion matrix and classification report in training set are:
[[ 24  11]
 [ 12 105]]
              precision    recall  f1-score   support

     Healthy       0.69      0.67      0.68        36
   Parkinson       0.90      0.91      0.90       116

    accuracy                           0.85       152
   macro avg       0.79      0.79      0.79       152
weighted avg       0.85      0.85      0.85       152

The accuracy in testing set is: 69.77%
The precision in testing set is: 71.43%
The recall in testing set is: 96.77%
The F1 score in testing set is: 82.19%
The parameters after the GridSearrch are:  {'var_smoothing': 1.873817422860383e-05}
The confusion matrix and classification report in training set are:
[[ 0  1]
 [12 30]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        12
           1       0.71      0.97      0.82        31

    accuracy                           0.70        43
   macro avg       0.36      0.48      0.41        43
weighted avg       0.51      0.70      0.59        43

################################################################################################################
########################
######                                  DECISION TREE MODEL (0ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
```

```
DecisionTreeClassifier(class_weight={0: 1, 1: 1}, max_depth=2, max_features=5) {'ccp_alpha': 0.
0, 'class_weight': {0: 1, 1: 1}, 'criterion': 'gini', 'max_depth': 2, 'max_features': 5}
The accuracy in training set is: 92.76%
The precision in training set is: 92.00%
The recall in training set is: 99.14%
The F1 score in training set is: 95.44%
The confusion matrix and classification report in training set are:
[[ 26   1]
 [ 10 115]]
              precision    recall  f1-score   support

     Healthy       0.96      0.72      0.83        36
   Parkinson       0.92      0.99      0.95       116

    accuracy                           0.93       152
   macro avg       0.94      0.86      0.89       152
weighted avg       0.93      0.93      0.92       152

The accuracy in testing set is: 72.09%
The precison in testing set is: 72.09%
The recall in testing set is: 100.00%
The F1 score in testing set is: 83.78%
The parameters after the GridSearrch are:  {'ccp_alpha': 0.0, 'class_weight': {0: 1, 1: 1}, 'cr
iterion': 'gini', 'max_depth': 2, 'max_features': 5}
The confusion matrix and classification report in training set are:
[[ 0  0]
 [12 31]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        12
           1       0.72      1.00      0.84        31

    accuracy                           0.72        43
   macro avg       0.36      0.50      0.42        43
weighted avg       0.52      0.72      0.60        43

##########################################################################################
#######################
######                                LOGISTIC REGRESSION MODEL (0ος διαχωρισμος)
######
######
######
After grid search the Logistic model has the follow parameters:
LogisticRegression(C=1, class_weight={0: 1, 1: 3}, tol=1e-08) {'C': 1, 'class_weight': {0: 1, 1
: 3}, 'max_iter': 100, 'penalty': 'l2', 'solver': 'lbfgs', 'tol': 1e-08}
The accuracy in training set is: 91.45%
The precion in training set is: 89.92%
The recall in training set is: 100.00%
The F1 score in training set is: 94.69%
```

```
The confusion matrix and classification report in training set are:
[[ 23   0]
 [ 13 116]]
              precision    recall  f1-score   support

     Healthy       1.00      0.64      0.78        36
   Parkinson       0.90      1.00      0.95       116

    accuracy                           0.91       152
   macro avg       0.95      0.82      0.86       152
weighted avg       0.92      0.91      0.91       152

The accuracy in testing set is: 72.09%
The precion testing set is: 72.09%
The recall in testing set is: 100.00%
The F1 score in testing set is: 83.78%
The parameters after the GridSearrch are:  {'C': 1, 'class_weight': {0: 1, 1: 3}, 'max_iter': 1
00, 'penalty': 'l2', 'solver': 'lbfgs', 'tol': 1e-08}
The confusion matrix and classification report in training set are:
[[ 0  0]
 [12 31]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        12
           1       0.72      1.00      0.84        31

    accuracy                           0.72        43
   macro avg       0.36      0.50      0.42        43
weighted avg       0.52      0.72      0.60        43

##########################################################################################
######################
######                              SUPPORT VECTOR MACHINE MODEL (0ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=26.826957952797247, class_weight='balanced', gamma=1e-09) {'C': 26.826957952797247, 'clas
s_weight': 'balanced', 'gamma': 1e-09, 'kernel': 'rbf'}
The accuracy in training set is: 23.68%
The precision in training set is: 0.00%
The recall in training set is: 0.00%
The F1 score in training set is: 0.00%
The confusion matrix and classification report in training set are:
[[ 36 116]
 [  0   0]]
              precision    recall  f1-score   support

     Healthy       0.24      1.00      0.38        36
```

```
   Parkinson       0.00      0.00      0.00       116

    accuracy                           0.24       152
   macro avg       0.12      0.50      0.19       152
weighted avg       0.06      0.24      0.09       152

The accuracy in testing set is: 27.91%
The precision in testing set is: 0.00%
The recall in testing set is: 0.00%
The F1 score in testing set is: 0.00%
The parameters after the GridSearrch are:  {'C': 26.826957952797247, 'class_weight': 'balanced'
, 'gamma': 1e-09, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[12 31]
 [ 0  0]]
              precision    recall  f1-score   support

           0       0.28      1.00      0.44        12
           1       0.00      0.00      0.00        31

    accuracy                           0.28        43
   macro avg       0.14      0.50      0.22        43
weighted avg       0.08      0.28      0.12        43

[03:59:13] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.

  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core.  Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.


[03:59:13] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
######                                          XGB MODEL (0ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.5, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
```

```
              monotone_constraints='()', n_estimators=5, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.3, tree_method='exact',
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.5, 'n_estimators': 5, 'subsample': 0.3}
The accuracy in training set is: 92.76%
The precision in training set is: 93.39%
The recall in training set is: 97.41%
The F1 score in training set is: 95.36%
The confusion matrix and classification report in training set are:
[[ 28   3]
 [  8 113]]
              precision    recall  f1-score   support

     Healthy       0.90      0.78      0.84        36
   Parkinson       0.93      0.97      0.95       116

    accuracy                           0.93       152
   macro avg       0.92      0.88      0.89       152
weighted avg       0.93      0.93      0.93       152

The accuracy in testing set is: 76.74%
The precision in testing set is: 75.61%
The recall in testing set is: 100.00%
The F1 score in testing set is: 86.11%
The parameters after the GridSearrch are:  {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.5
, 'n_estimators': 5, 'subsample': 0.3}
The confusion matrix and classification report in training set are:
[[ 2  0]
 [10 31]]
              precision    recall  f1-score   support

           0       1.00      0.17      0.29        12
           1       0.76      1.00      0.86        31

    accuracy                           0.77        43
   macro avg       0.88      0.58      0.57        43
weighted avg       0.82      0.77      0.70        43

##########################################################################################
######################
######                                    RANDOM FOREST MODEL (0ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(class_weight='balanced', max_depth=2,
```

min_samples_split=5, oob_score=True) {'bootstrap': True, 'class_weight': 'balanced', 'criterion': 'gini', 'max_depth': 2, 'max_features': 'auto', 'min_samples_split': 5, 'n_estimators': 100, 'oob_score': True}
The accuracy in training set is: 94.08%
The precision in training set is: 94.96%
The recall in training set is: 97.41%
The F1 score in training set is: 96.17%
The confusion matrix and classification report in training set are:
[[30   3]
 [6 113]]

	precision	recall	f1-score	support
Healthy	0.91	0.83	0.87	36
Parkinson	0.95	0.97	0.96	116
accuracy			0.94	152
macro avg	0.93	0.90	0.92	152
weighted avg	0.94	0.94	0.94	152

The accuracy in testing set is: 69.77%
The precision in testing set is: 73.68%
The recall in testing set is: 90.32%
The F1 score in testing set is: 81.16%
The parameters after the GridSearrch are: {'bootstrap': True, 'class_weight': 'balanced', 'criterion': 'gini', 'max_depth': 2, 'max_features': 'auto', 'min_samples_split': 5, 'n_estimators': 100, 'oob_score': True}
The confusion matrix and classification report in training set are:
[[2 3]
 [10 28]]

	precision	recall	f1-score	support
0	0.40	0.17	0.24	12
1	0.74	0.90	0.81	31
accuracy			0.70	43
macro avg	0.57	0.53	0.52	43
weighted avg	0.64	0.70	0.65	43

Στον 1ο διαχωρισμό του dataset έχουμε:
##
######################
GAUSSIAN_NB MODEL (1ος διαχωρισμος)
######
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=1.0) {'var_smoothing': 1.0}
The accuracy in training set is: 79.74%
The precision in training set is: 80.67%

```
The recall in training set is: 98.37%
The F1 score in training set is: 88.64%
The confusion matrix and classification report in training set are:
[[  1   2]
 [ 29 121]]
              precision    recall  f1-score   support

     Healthy       0.33      0.03      0.06        30
   Parkinson       0.81      0.98      0.89       123

    accuracy                           0.80       153
   macro avg       0.57      0.51      0.47       153
weighted avg       0.71      0.80      0.72       153

The accuracy in testing set is: 54.76%
The precision in testing set is: 56.10%
The recall in testing set is: 95.83%
The F1 score in testing set is: 70.77%
The parameters after the GridSearrch are:  {'var_smoothing': 1.0}
The confusion matrix and classification report in training set are:
[[ 0  1]
 [18 23]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        18
           1       0.56      0.96      0.71        24

    accuracy                           0.55        42
   macro avg       0.28      0.48      0.35        42
weighted avg       0.32      0.55      0.40        42

##########################################################################################
#######################
######                                DECISION TREE MODEL (1ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
DecisionTreeClassifier(ccp_alpha=0.2, class_weight='balanced', max_depth=2,
                       max_features=1) {'ccp_alpha': 0.2, 'class_weight': 'balanced', 'criterio
n': 'gini', 'max_depth': 2, 'max_features': 1}
The accuracy in training set is: 80.39%
The precision in training set is: 80.39%
The recall in training set is: 100.00%
The F1 score in training set is: 89.13%
The confusion matrix and classification report in training set are:
[[  0   0]
 [ 30 123]]
              precision    recall  f1-score   support
```

```
     Healthy       0.00      0.00      0.00        30
   Parkinson       0.80      1.00      0.89       123

    accuracy                           0.80       153
   macro avg       0.40      0.50      0.45       153
weighted avg       0.65      0.80      0.72       153

The accuracy in testing set is: 57.14%
The precison in testing set is: 57.14%
The recall in testing set is: 100.00%
The F1 score in testing set is: 72.73%
The parameters after the GridSearrch are:  {'ccp_alpha': 0.2, 'class_weight': 'balanced', 'crit
erion': 'gini', 'max_depth': 2, 'max_features': 1}
The confusion matrix and classification report in training set are:
[[ 0  0]
 [18 24]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        18
           1       0.57      1.00      0.73        24

    accuracy                           0.57        42
   macro avg       0.29      0.50      0.36        42
weighted avg       0.33      0.57      0.42        42

##########################################################################################
#######################
######                                LOGISTIC REGRESSION MODEL (1ος διαχωρισμος)
######
######
######
After grid search the Logistic model has the follow parameters:
LogisticRegression(C=0.0001, class_weight={0: 1, 1: 1}, penalty='l1',
                   solver='saga', tol=1e-08) {'C': 0.0001, 'class_weight': {0: 1, 1: 1}, 'max_i
ter': 100, 'penalty': 'l1', 'solver': 'saga', 'tol': 1e-08}
The accuracy in training set is: 80.39%
The precion in training set is: 80.39%
The recall in training set is: 100.00%
The F1 score in training set is: 89.13%
The confusion matrix and classification report in training set are:
[[  0   0]
 [ 30 123]]
              precision    recall  f1-score   support

     Healthy       0.00      0.00      0.00        30
   Parkinson       0.80      1.00      0.89       123

    accuracy                           0.80       153
```

```
   macro avg       0.40      0.50      0.45       153
weighted avg       0.65      0.80      0.72       153

The accuracy in testing set is: 57.14%
The precion testing set is: 57.14%
The recall in testing set is: 100.00%
The F1 score in testing set is: 72.73%
The parameters after the GridSearrch are:  {'C': 0.0001, 'class_weight': {0: 1, 1: 1}, 'max_ite
r': 100, 'penalty': 'l1', 'solver': 'saga', 'tol': 1e-08}
The confusion matrix and classification report in training set are:
[[ 0  0]
 [18 24]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        18
           1       0.57      1.00      0.73        24

    accuracy                           0.57        42
   macro avg       0.29      0.50      0.36        42
weighted avg       0.33      0.57      0.42        42

##########################################################################################
#######################
######                            SUPPORT VECTOR MACHINE MODEL (1ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=3727593.720314938, class_weight={0: 1, 1: 3}, gamma=1e-05) {'C': 3727593.720314938, 'clas
s_weight': {0: 1, 1: 3}, 'gamma': 1e-05, 'kernel': 'rbf'}
The accuracy in training set is: 99.35%
The precision in training set is: 99.19%
The recall in training set is: 100.00%
The F1 score in training set is: 99.60%
The confusion matrix and classification report in training set are:
[[ 29   0]
 [  1 123]]
              precision    recall  f1-score   support

     Healthy       1.00      0.97      0.98        30
   Parkinson       0.99      1.00      1.00       123

    accuracy                           0.99       153
   macro avg       1.00      0.98      0.99       153
weighted avg       0.99      0.99      0.99       153

The accuracy in testing set is: 57.14%
The precision in testing set is: 61.54%
The recall in testing set is: 66.67%
```

```
The F1 score in testing set is: 64.00%
The parameters after the GridSearrch are:  {'C': 3727593.720314938, 'class_weight': {0: 1, 1: 3
}, 'gamma': 1e-05, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[ 8  8]
 [10 16]]
              precision    recall  f1-score   support

           0       0.50      0.44      0.47        18
           1       0.62      0.67      0.64        24

    accuracy                           0.57        42
   macro avg       0.56      0.56      0.56        42
weighted avg       0.57      0.57      0.57        42

[04:31:04] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.

  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core.  Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.


[04:31:04] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
######                                          XGB MODEL (1ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.9, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
              monotone_constraints='()', n_estimators=5, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.5, tree_method='exact',
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.9, 'n_estimators': 5, 'subsample': 0.5}
The accuracy in training set is: 97.39%
The precision in training set is: 96.85%
The recall in training set is: 100.00%
The F1 score in training set is: 98.40%
```

```
The confusion matrix and classification report in training set are:
[[ 26   0]
 [  4 123]]
              precision    recall  f1-score   support

     Healthy       1.00      0.87      0.93        30
   Parkinson       0.97      1.00      0.98       123

    accuracy                           0.97       153
   macro avg       0.98      0.93      0.96       153
weighted avg       0.97      0.97      0.97       153

The accuracy in testing set is: 69.05%
The precision in testing set is: 70.37%
The recall in testing set is: 79.17%
The F1 score in testing set is: 74.51%
The parameters after the GridSearrch are:  {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.9
, 'n_estimators': 5, 'subsample': 0.5}
The confusion matrix and classification report in training set are:
[[10  5]
 [ 8 19]]
              precision    recall  f1-score   support

           0       0.67      0.56      0.61        18
           1       0.70      0.79      0.75        24

    accuracy                           0.69        42
   macro avg       0.69      0.67      0.68        42
weighted avg       0.69      0.69      0.69        42

##########################################################################################
######################
######                                        RANDOM FOREST MODEL (1ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(class_weight={0: 1, 1: 10}, max_depth=5,
                       min_samples_split=5, n_estimators=500, oob_score=True) {'bootstrap': Tru
e, 'class_weight': {0: 1, 1: 10}, 'criterion': 'gini', 'max_depth': 5, 'max_features': 'auto',
'min_samples_split': 5, 'n_estimators': 500, 'oob_score': True}
The accuracy in training set is: 92.16%
The precision in training set is: 91.11%
The recall in training set is: 100.00%
The F1 score in training set is: 95.35%
The confusion matrix and classification report in training set are:
[[ 18   0]
 [ 12 123]]
              precision    recall  f1-score   support
```

```
     Healthy       1.00      0.60      0.75        30
   Parkinson       0.91      1.00      0.95       123

    accuracy                           0.92       153
   macro avg       0.96      0.80      0.85       153
weighted avg       0.93      0.92      0.91       153

The accuracy in testing set is: 83.33%
The precision in testing set is: 77.42%
The recall in testing set is: 100.00%
The F1 score in testing set is: 87.27%
The parameters after the GridSearrch are:  {'bootstrap': True, 'class_weight': {0: 1, 1: 10}, '
criterion': 'gini', 'max_depth': 5, 'max_features': 'auto', 'min_samples_split': 5, 'n_estimato
rs': 500, 'oob_score': True}
The confusion matrix and classification report in training set are:
[[11  0]
 [ 7 24]]
              precision    recall  f1-score   support

           0       1.00      0.61      0.76        18
           1       0.77      1.00      0.87        24

    accuracy                           0.83        42
   macro avg       0.89      0.81      0.82        42
weighted avg       0.87      0.83      0.82        42

Στον 2ο διαχωρισμό του dataset έχουμε:
##############################################################################################
######################
######                                GAUSSIAN_NB MODEL (2ος διαχωρισμος)
######
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=0.15199110829529336) {'var_smoothing': 0.15199110829529336}
The accuracy in training set is: 94.12%
The precision in training set is: 94.12%
The recall in training set is: 99.22%
The F1 score in training set is: 96.60%
The confusion matrix and classification report in training set are:
[[ 16   1]
 [  8 128]]
              precision    recall  f1-score   support

     Healthy       0.94      0.67      0.78        24
   Parkinson       0.94      0.99      0.97       129

    accuracy                           0.94       153
```

```
   macro avg       0.94      0.83      0.87       153
weighted avg       0.94      0.94      0.94       153

The accuracy in testing set is: 42.86%
The precision in testing set is: 41.67%
The recall in testing set is: 83.33%
The F1 score in testing set is: 55.56%
The parameters after the GridSearrch are:  {'var_smoothing': 0.15199110829529336}
The confusion matrix and classification report in training set are:
[[ 3  3]
 [21 15]]
              precision    recall  f1-score   support

           0       0.50      0.12      0.20        24
           1       0.42      0.83      0.56        18

    accuracy                           0.43        42
   macro avg       0.46      0.48      0.38        42
weighted avg       0.46      0.43      0.35        42

##########################################################################################
######################
######                                DECISION TREE MODEL (2ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
DecisionTreeClassifier(class_weight={0: 1, 1: 10}, max_depth=2, max_features=5) {'ccp_alpha': 0
.0, 'class_weight': {0: 1, 1: 10}, 'criterion': 'gini', 'max_depth': 2, 'max_features': 5}
The accuracy in training set is: 94.77%
The precision in training set is: 94.16%
The recall in training set is: 100.00%
The F1 score in training set is: 96.99%
The confusion matrix and classification report in training set are:
[[ 16   0]
 [  8 129]]
              precision    recall  f1-score   support

     Healthy       1.00      0.67      0.80        24
   Parkinson       0.94      1.00      0.97       129

    accuracy                           0.95       153
   macro avg       0.97      0.83      0.88       153
weighted avg       0.95      0.95      0.94       153

The accuracy in testing set is: 52.38%
The precison in testing set is: 47.37%
The recall in testing set is: 100.00%
The F1 score in testing set is: 64.29%
```

```
The parameters after the GridSearrch are:  {'ccp_alpha': 0.0, 'class_weight': {0: 1, 1: 10}, 'c
riterion': 'gini', 'max_depth': 2, 'max_features': 5}
The confusion matrix and classification report in training set are:
[[ 4  0]
 [20 18]]
              precision    recall  f1-score   support

           0       1.00      0.17      0.29        24
           1       0.47      1.00      0.64        18

    accuracy                           0.52        42
   macro avg       0.74      0.58      0.46        42
weighted avg       0.77      0.52      0.44        42

##########################################################################################
#######################
######                                LOGISTIC REGRESSION MODEL (2ος διαχωρισμος)
######
######
######
After grid search the Logistic model has the follow parameters:
LogisticRegression(C=1, class_weight={0: 1, 1: 3}, penalty='l1',
                   solver='liblinear', tol=0.01) {'C': 1, 'class_weight': {0: 1, 1: 3}, 'max_it
er': 100, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The accuracy in training set is: 94.12%
The precion in training set is: 93.48%
The recall in training set is: 100.00%
The F1 score in training set is: 96.63%
The confusion matrix and classification report in training set are:
[[ 15   0]
 [  9 129]]
              precision    recall  f1-score   support

     Healthy       1.00      0.62      0.77        24
   Parkinson       0.93      1.00      0.97       129

    accuracy                           0.94       153
   macro avg       0.97      0.81      0.87       153
weighted avg       0.95      0.94      0.94       153

The accuracy in testing set is: 50.00%
The precion testing set is: 46.15%
The recall in testing set is: 100.00%
The F1 score in testing set is: 63.16%
The parameters after the GridSearrch are:  {'C': 1, 'class_weight': {0: 1, 1: 3}, 'max_iter': 1
00, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The confusion matrix and classification report in training set are:
[[ 3  0]
 [21 18]]
```

```
              precision    recall  f1-score   support

           0       1.00      0.12      0.22        24
           1       0.46      1.00      0.63        18

    accuracy                           0.50        42
   macro avg       0.73      0.56      0.43        42
weighted avg       0.77      0.50      0.40        42

################################################################################################
######################
######                              SUPPORT VECTOR MACHINE MODEL (2ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=10000.0, class_weight={0: 1, 1: 1}, gamma=2.1544346900318822e-08) {'C': 10000.0, 'class_w
eight': {0: 1, 1: 1}, 'gamma': 2.1544346900318822e-08, 'kernel': 'rbf'}
The accuracy in training set is: 94.12%
The precision in training set is: 93.48%
The recall in training set is: 100.00%
The F1 score in training set is: 96.63%
The confusion matrix and classification report in training set are:
[[ 15   0]
 [  9 129]]
              precision    recall  f1-score   support

     Healthy       1.00      0.62      0.77        24
   Parkinson       0.93      1.00      0.97       129

    accuracy                           0.94       153
   macro avg       0.97      0.81      0.87       153
weighted avg       0.95      0.94      0.94       153

The accuracy in testing set is: 47.62%
The precision in testing set is: 44.74%
The recall in testing set is: 94.44%
The F1 score in testing set is: 60.71%
The parameters after the GridSearrch are:  {'C': 10000.0, 'class_weight': {0: 1, 1: 1}, 'gamma'
: 2.1544346900318822e-08, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[ 3  1]
 [21 17]]
              precision    recall  f1-score   support

           0       0.75      0.12      0.21        24
           1       0.45      0.94      0.61        18

    accuracy                           0.48        42
```

```
   macro avg       0.60      0.53      0.41        42
weighted avg       0.62      0.48      0.38        42

[05:02:52] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.

  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core.  Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.


[05:02:52] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
######                                          XGB MODEL (2ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.5, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
              monotone_constraints='()', n_estimators=150, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.5, tree_method='exact',
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.5, 'n_estimators': 150, 'subsample': 0.5}
The accuracy in training set is: 99.35%
The precision in training set is: 99.23%
The recall in training set is: 100.00%
The F1 score in training set is: 99.61%
The confusion matrix and classification report in training set are:
[[ 23   0]
 [  1 129]]
              precision    recall  f1-score   support

     Healthy       1.00      0.96      0.98        24
   Parkinson       0.99      1.00      1.00       129

    accuracy                           0.99       153
   macro avg       1.00      0.98      0.99       153
weighted avg       0.99      0.99      0.99       153
```

```
The accuracy in testing set is: 57.14%
The precision in testing set is: 50.00%
The recall in testing set is: 88.89%
The F1 score in testing set is: 64.00%
The parameters after the GridSearrch are:  {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.5
, 'n_estimators': 150, 'subsample': 0.5}
The confusion matrix and classification report in training set are:
[[ 8  2]
 [16 16]]
              precision    recall  f1-score   support

           0       0.80      0.33      0.47        24
           1       0.50      0.89      0.64        18

    accuracy                           0.57        42
   macro avg       0.65      0.61      0.56        42
weighted avg       0.67      0.57      0.54        42

##########################################################################################
########################
######                                        RANDOM FOREST MODEL (2ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(class_weight={0: 1, 1: 3}, max_depth=10,
                       min_samples_split=5) {'bootstrap': True, 'class_weight': {0: 1, 1: 3}, '
criterion': 'gini', 'max_depth': 10, 'max_features': 'auto', 'min_samples_split': 5, 'n_estimat
ors': 100, 'oob_score': False}
The accuracy in training set is: 100.00%
The precision in training set is: 100.00%
The recall in training set is: 100.00%
The F1 score in training set is: 100.00%
The confusion matrix and classification report in training set are:
[[ 24   0]
 [  0 129]]
              precision    recall  f1-score   support

     Healthy       1.00      1.00      1.00        24
   Parkinson       1.00      1.00      1.00       129

    accuracy                           1.00       153
   macro avg       1.00      1.00      1.00       153
weighted avg       1.00      1.00      1.00       153

The accuracy in testing set is: 47.62%
The precision in testing set is: 44.44%
The recall in testing set is: 88.89%
The F1 score in testing set is: 59.26%
```

```
The parameters after the GridSearrch are:  {'bootstrap': True, 'class_weight': {0: 1, 1: 3}, 'c
riterion': 'gini', 'max_depth': 10, 'max_features': 'auto', 'min_samples_split': 5, 'n_estimato
rs': 100, 'oob_score': False}
The confusion matrix and classification report in training set are:
[[ 4  2]
 [20 16]]
              precision    recall  f1-score   support

           0       0.67      0.17      0.27        24
           1       0.44      0.89      0.59        18

    accuracy                           0.48        42
   macro avg       0.56      0.53      0.43        42
weighted avg       0.57      0.48      0.41        42

Στον 3ο διαχωρισμό του dataset έχουμε:
################################################################################################
######################
######                                GAUSSIAN_NB MODEL (3ος διαχωρισμος)
######
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=1.2328467394420658e-05) {'var_smoothing': 1.2328467394420658e-05}
The accuracy in training set is: 81.70%
The precision in training set is: 83.20%
The recall in training set is: 93.69%
The F1 score in training set is: 88.14%
The confusion matrix and classification report in training set are:
[[ 21   7]
 [ 21 104]]
              precision    recall  f1-score   support

     Healthy       0.75      0.50      0.60        42
   Parkinson       0.83      0.94      0.88       111

    accuracy                           0.82       153
   macro avg       0.79      0.72      0.74       153
weighted avg       0.81      0.82      0.80       153

The accuracy in testing set is: 92.86%
The precision in testing set is: 100.00%
The recall in testing set is: 91.67%
The F1 score in testing set is: 95.65%
The parameters after the GridSearrch are:  {'var_smoothing': 1.2328467394420658e-05}
The confusion matrix and classification report in training set are:
[[ 6  3]
 [ 0 33]]
              precision    recall  f1-score   support
```

```
           0       0.67      1.00      0.80         6
           1       1.00      0.92      0.96        36

    accuracy                           0.93        42
   macro avg       0.83      0.96      0.88        42
weighted avg       0.95      0.93      0.93        42

##########################################################################################
######################
######                                DECISION TREE MODEL (3ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
DecisionTreeClassifier(ccp_alpha=0.2, class_weight='balanced', max_depth=50,
                       max_features=1) {'ccp_alpha': 0.2, 'class_weight': 'balanced', 'criterio
n': 'gini', 'max_depth': 50, 'max_features': 1}
The accuracy in training set is: 27.45%
The precision in training set is: 0.00%
The recall in training set is: 0.00%
The F1 score in training set is: 0.00%
The confusion matrix and classification report in training set are:
[[ 42 111]
 [  0   0]]
              precision    recall  f1-score   support

     Healthy       0.27      1.00      0.43        42
   Parkinson       0.00      0.00      0.00       111

    accuracy                           0.27       153
   macro avg       0.14      0.50      0.22       153
weighted avg       0.08      0.27      0.12       153

The accuracy in testing set is: 14.29%
The precison in testing set is: 0.00%
The recall in testing set is: 0.00%
The F1 score in testing set is: 0.00%
The parameters after the GridSearrch are:  {'ccp_alpha': 0.2, 'class_weight': 'balanced', 'crit
erion': 'gini', 'max_depth': 50, 'max_features': 1}
The confusion matrix and classification report in training set are:
[[ 6 36]
 [ 0  0]]
              precision    recall  f1-score   support

           0       0.14      1.00      0.25         6
           1       0.00      0.00      0.00        36

    accuracy                           0.14        42
```

```
   macro avg       0.07      0.50      0.12        42
weighted avg       0.02      0.14      0.04        42

##########################################################################################
######################
######                                LOGISTIC REGRESSION MODEL (3ος διαχωρισμος)
######
######
######
After grid search the Logistic model has the follow parameters:
LogisticRegression(C=1, class_weight={0: 1, 1: 1}, solver='liblinear', tol=0.01) {'C': 1, 'clas
s_weight': {0: 1, 1: 1}, 'max_iter': 100, 'penalty': 'l2', 'solver': 'liblinear', 'tol': 0.01}
The accuracy in training set is: 85.62%
The precion in training set is: 85.60%
The recall in training set is: 96.40%
The F1 score in training set is: 90.68%
The confusion matrix and classification report in training set are:
[[ 24   4]
 [ 18 107]]
              precision    recall  f1-score   support

     Healthy       0.86      0.57      0.69        42
   Parkinson       0.86      0.96      0.91       111

    accuracy                           0.86       153
   macro avg       0.86      0.77      0.80       153
weighted avg       0.86      0.86      0.85       153

The accuracy in testing set is: 80.95%
The precion testing set is: 100.00%
The recall in testing set is: 77.78%
The F1 score in testing set is: 87.50%
The parameters after the GridSearrch are:  {'C': 1, 'class_weight': {0: 1, 1: 1}, 'max_iter': 1
00, 'penalty': 'l2', 'solver': 'liblinear', 'tol': 0.01}
The confusion matrix and classification report in training set are:
[[ 6  8]
 [ 0 28]]
              precision    recall  f1-score   support

           0       0.43      1.00      0.60         6
           1       1.00      0.78      0.88        36

    accuracy                           0.81        42
   macro avg       0.71      0.89      0.74        42
weighted avg       0.92      0.81      0.84        42

##########################################################################################
######################
```

SUPPORT VECTOR MACHINE MODEL (3ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=0.07196856730011521, class_weight='balanced', gamma=2.154434690031878) {'C': 0.0719685673
0011521, 'class_weight': 'balanced', 'gamma': 2.154434690031878, 'kernel': 'rbf'}
The accuracy in training set is: 72.55%
The precision in training set is: 72.55%
The recall in training set is: 100.00%
The F1 score in training set is: 84.09%
The confusion matrix and classification report in training set are:
[[0   0]
 [42 111]]
              precision   recall f1-score   support

     Healthy   0.00   0.00   0.00   42
   Parkinson   0.73   1.00   0.84   111

    accuracy   0.73   153
   macro avg   0.36   0.50   0.42   153
weighted avg   0.53   0.73   0.61   153

The accuracy in testing set is: 85.71%
The precision in testing set is: 85.71%
The recall in testing set is: 100.00%
The F1 score in testing set is: 92.31%
The parameters after the GridSearrch are: {'C': 0.07196856730011521, 'class_weight': 'balanced
', 'gamma': 2.154434690031878, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[0 0]
 [6 36]]
              precision   recall f1-score   support

           0   0.00   0.00   0.00   6
           1   0.86   1.00   0.92   36

    accuracy   0.86   42
   macro avg   0.43   0.50   0.46   42
weighted avg   0.73   0.86   0.79   42

[05:36:17] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.

  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core. Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.

```
[05:36:17] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
######                                          XGB MODEL (3ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.1, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
              monotone_constraints='()', n_estimators=5, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.5, tree_method='exact',
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.1, 'n_estimators': 5, 'subsample': 0.5}
The accuracy in training set is: 90.85%
The precision in training set is: 90.76%
The recall in training set is: 97.30%
The F1 score in training set is: 93.91%
The confusion matrix and classification report in training set are:
[[ 31   3]
 [ 11 108]]
              precision    recall  f1-score   support

     Healthy       0.91      0.74      0.82        42
   Parkinson       0.91      0.97      0.94       111

    accuracy                           0.91       153
   macro avg       0.91      0.86      0.88       153
weighted avg       0.91      0.91      0.91       153

The accuracy in testing set is: 88.10%
The precision in testing set is: 100.00%
The recall in testing set is: 86.11%
The F1 score in testing set is: 92.54%
The parameters after the GridSearrch are:  {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.1
, 'n_estimators': 5, 'subsample': 0.5}
The confusion matrix and classification report in training set are:
[[ 6  5]
 [ 0 31]]
              precision    recall  f1-score   support
```

```
           0       0.55      1.00      0.71         6
           1       1.00      0.86      0.93        36

    accuracy                           0.88        42
   macro avg       0.77      0.93      0.82        42
weighted avg       0.94      0.88      0.89        42

################################################################################################
######################
######                                        RANDOM FOREST MODEL (3ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(class_weight='balanced', max_depth=10,
                       max_features='sqrt', min_samples_split=10) {'bootstrap': True, 'class_we
ight': 'balanced', 'criterion': 'gini', 'max_depth': 10, 'max_features': 'sqrt', 'min_samples_s
plit': 10, 'n_estimators': 100, 'oob_score': False}
The accuracy in training set is: 98.69%
The precision in training set is: 100.00%
The recall in training set is: 98.20%
The F1 score in training set is: 99.09%
The confusion matrix and classification report in training set are:
[[ 42   2]
 [  0 109]]
              precision    recall  f1-score   support

     Healthy       0.95      1.00      0.98        42
   Parkinson       1.00      0.98      0.99       111

    accuracy                           0.99       153
   macro avg       0.98      0.99      0.98       153
weighted avg       0.99      0.99      0.99       153

The accuracy in testing set is: 80.95%
The precision in testing set is: 93.75%
The recall in testing set is: 83.33%
The F1 score in testing set is: 88.24%
The parameters after the GridSearrch are:  {'bootstrap': True, 'class_weight': 'balanced', 'cri
terion': 'gini', 'max_depth': 10, 'max_features': 'sqrt', 'min_samples_split': 10, 'n_estimator
s': 100, 'oob_score': False}
The confusion matrix and classification report in training set are:
[[ 4  6]
 [ 2 30]]
              precision    recall  f1-score   support

           0       0.40      0.67      0.50         6
           1       0.94      0.83      0.88        36
```

```
    accuracy                           0.81        42
   macro avg       0.67      0.75      0.69        42
weighted avg       0.86      0.81      0.83        42

Στον 4ο διαχωρισμό του dataset έχουμε:
################################################################################################
########################
######                                  GAUSSIAN_NB MODEL (4ος διαχωρισμος)
######
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=0.2848035868435802) {'var_smoothing': 0.2848035868435802}
The accuracy in training set is: 80.92%
The precision in training set is: 80.00%
The recall in training set is: 98.18%
The F1 score in training set is: 88.16%
The confusion matrix and classification report in training set are:
[[ 15   2]
 [ 27 108]]
              precision    recall  f1-score   support

     Healthy       0.88      0.36      0.51        42
   Parkinson       0.80      0.98      0.88       110

    accuracy                           0.81       152
   macro avg       0.84      0.67      0.70       152
weighted avg       0.82      0.81      0.78       152

The accuracy in testing set is: 83.72%
The precision in testing set is: 87.50%
The recall in testing set is: 94.59%
The F1 score in testing set is: 90.91%
The parameters after the GridSearrch are:  {'var_smoothing': 0.2848035868435802}
The confusion matrix and classification report in training set are:
[[ 1  2]
 [ 5 35]]
              precision    recall  f1-score   support

           0       0.33      0.17      0.22         6
           1       0.88      0.95      0.91        37

    accuracy                           0.84        43
   macro avg       0.60      0.56      0.57        43
weighted avg       0.80      0.84      0.81        43

################################################################################################
########################
```

```
######                                DECISION TREE MODEL (4ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
DecisionTreeClassifier(ccp_alpha=0.4, class_weight='balanced', max_depth=2,
                       max_features=1) {'ccp_alpha': 0.4, 'class_weight': 'balanced', 'criterio
n': 'gini', 'max_depth': 2, 'max_features': 1}
The accuracy in training set is: 27.63%
The precision in training set is: 0.00%
The recall in training set is: 0.00%
The F1 score in training set is: 0.00%
The confusion matrix and classification report in training set are:
[[ 42 110]
 [  0   0]]
              precision    recall  f1-score   support

     Healthy       0.28      1.00      0.43        42
   Parkinson       0.00      0.00      0.00       110

    accuracy                           0.28       152
   macro avg       0.14      0.50      0.22       152
weighted avg       0.08      0.28      0.12       152

The accuracy in testing set is: 13.95%
The precison in testing set is: 0.00%
The recall in testing set is: 0.00%
The F1 score in testing set is: 0.00%
The parameters after the GridSearrch are:  {'ccp_alpha': 0.4, 'class_weight': 'balanced', 'crit
erion': 'gini', 'max_depth': 2, 'max_features': 1}
The confusion matrix and classification report in training set are:
[[ 6 37]
 [ 0  0]]
              precision    recall  f1-score   support

           0       0.14      1.00      0.24         6
           1       0.00      0.00      0.00        37

    accuracy                           0.14        43
   macro avg       0.07      0.50      0.12        43
weighted avg       0.02      0.14      0.03        43

#########################################################################################
######################
######                                LOGISTIC REGRESSION MODEL (4ος διαχωρισμος)
######
######
######
After grid search the Logistic model has the follow parameters:
```

```
LogisticRegression(C=1, class_weight={0: 1, 1: 3}, max_iter=1000, penalty='l1',
                   solver='liblinear', tol=0.01) {'C': 1, 'class_weight': {0: 1, 1: 3}, 'max_it
er': 1000, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The accuracy in training set is: 82.24%
The precion in training set is: 80.29%
The recall in training set is: 100.00%
The F1 score in training set is: 89.07%
The confusion matrix and classification report in training set are:
[[ 15   0]
 [ 27 110]]
              precision    recall  f1-score   support

     Healthy       1.00      0.36      0.53        42
   Parkinson       0.80      1.00      0.89       110

    accuracy                           0.82       152
   macro avg       0.90      0.68      0.71       152
weighted avg       0.86      0.82      0.79       152

The accuracy in testing set is: 100.00%
The precion testing set is: 100.00%
The recall in testing set is: 100.00%
The F1 score in testing set is: 100.00%
The parameters after the GridSearrch are:  {'C': 1, 'class_weight': {0: 1, 1: 3}, 'max_iter': 1
000, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The confusion matrix and classification report in training set are:
[[ 6  0]
 [ 0 37]]
              precision    recall  f1-score   support

           0       1.00      1.00      1.00         6
           1       1.00      1.00      1.00        37

    accuracy                           1.00        43
   macro avg       1.00      1.00      1.00        43
weighted avg       1.00      1.00      1.00        43

##########################################################################################
########################
######                            SUPPORT VECTOR MACHINE MODEL (4ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=26826957.95279727, class_weight={0: 1, 1: 3},
    gamma=2.1544346900318822e-08) {'C': 26826957.95279727, 'class_weight': {0: 1, 1: 3}, 'gamma
': 2.1544346900318822e-08, 'kernel': 'rbf'}
The accuracy in training set is: 84.21%
The precision in training set is: 82.58%
```

```
The recall in training set is: 99.09%
The F1 score in training set is: 90.08%
The confusion matrix and classification report in training set are:
[[ 19   1]
 [ 23 109]]
              precision    recall  f1-score   support

     Healthy       0.95      0.45      0.61        42
   Parkinson       0.83      0.99      0.90       110

    accuracy                           0.84       152
   macro avg       0.89      0.72      0.76       152
weighted avg       0.86      0.84      0.82       152

The accuracy in testing set is: 97.67%
The precision in testing set is: 100.00%
The recall in testing set is: 97.30%
The F1 score in testing set is: 98.63%
The parameters after the GridSearrch are:  {'C': 26826957.95279727, 'class_weight': {0: 1, 1: 3
}, 'gamma': 2.1544346900318822e-08, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[ 6  1]
 [ 0 36]]
              precision    recall  f1-score   support

           0       0.86      1.00      0.92         6
           1       1.00      0.97      0.99        37

    accuracy                           0.98        43
   macro avg       0.93      0.99      0.95        43
weighted avg       0.98      0.98      0.98        43

[06:09:21] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.

  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core.  Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.


[06:09:21] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
######                                        XGB MODEL (4ος διαχωρισμος)
######
```

```
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.01, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
              monotone_constraints='()', n_estimators=5, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.5, tree_method='exact',
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.01, 'n_estimators': 5, 'subsample': 0.5}
The accuracy in training set is: 90.79%
The precision in training set is: 91.38%
The recall in training set is: 96.36%
The F1 score in training set is: 93.81%
The confusion matrix and classification report in training set are:
[[ 32   4]
 [ 10 106]]
              precision    recall  f1-score   support

     Healthy       0.89      0.76      0.82        42
   Parkinson       0.91      0.96      0.94       110

    accuracy                           0.91       152
   macro avg       0.90      0.86      0.88       152
weighted avg       0.91      0.91      0.91       152

The accuracy in testing set is: 86.05%
The precision in testing set is: 100.00%
The recall in testing set is: 83.78%
The F1 score in testing set is: 91.18%
The parameters after the GridSearrch are:  {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.0
1, 'n_estimators': 5, 'subsample': 0.5}
The confusion matrix and classification report in training set are:
[[ 6  6]
 [ 0 31]]
              precision    recall  f1-score   support

           0       0.50      1.00      0.67         6
           1       1.00      0.84      0.91        37

    accuracy                           0.86        43
   macro avg       0.75      0.92      0.79        43
weighted avg       0.93      0.86      0.88        43

##########################################################################################
######################
```

RANDOM FOREST MODEL (4ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(class_weight={0: 1, 1: 1}, max_depth=2,
                       max_features='sqrt', n_estimators=300, oob_score=True) {'bootstrap': Tru
e, 'class_weight': {0: 1, 1: 1}, 'criterion': 'gini', 'max_depth': 2, 'max_features': 'sqrt', '
min_samples_split': 2, 'n_estimators': 300, 'oob_score': True}
The accuracy in training set is: 88.82%
The precision in training set is: 86.61%
The recall in training set is: 100.00%
The F1 score in training set is: 92.83%
The confusion matrix and classification report in training set are:
[[25   0]
 [17 110]]
              precision   recall f1-score   support

     Healthy   1.00   0.60   0.75   42
   Parkinson   0.87   1.00   0.93   110

    accuracy   0.89   152
   macro avg   0.93   0.80   0.84   152
weighted avg   0.90   0.89   0.88   152

The accuracy in testing set is: 90.70%
The precision in testing set is: 100.00%
The recall in testing set is: 89.19%
The F1 score in testing set is: 94.29%
The parameters after the GridSearrch are: {'bootstrap': True, 'class_weight': {0: 1, 1: 1}, 'c
riterion': 'gini', 'max_depth': 2, 'max_features': 'sqrt', 'min_samples_split': 2, 'n_estimator
s': 300, 'oob_score': True}
The confusion matrix and classification report in training set are:
[[6 4]
 [0 33]]
              precision   recall f1-score   support

           0   0.60   1.00   0.75   6
           1   1.00   0.89   0.94   37

    accuracy   0.91   43
   macro avg   0.80   0.95   0.85   43
weighted avg   0.94   0.91   0.92   43

Στον 5ο διαχωρισμό του dataset έχουμε:
##
######################
GAUSSIAN_NB MODEL (5ος διαχωρισμος)
######

```
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=1.519911082952933e-06) {'var_smoothing': 1.519911082952933e-06}
The accuracy in training set is: 84.31%
The precision in training set is: 87.80%
The recall in training set is: 92.31%
The F1 score in training set is: 90.00%
The confusion matrix and classification report in training set are:
[[ 21   9]
 [ 15 108]]
              precision    recall  f1-score   support

     Healthy       0.70      0.58      0.64        36
   Parkinson       0.88      0.92      0.90       117

    accuracy                           0.84       153
   macro avg       0.79      0.75      0.77       153
weighted avg       0.84      0.84      0.84       153

The accuracy in testing set is: 88.10%
The precision in testing set is: 87.88%
The recall in testing set is: 96.67%
The F1 score in testing set is: 92.06%
The parameters after the GridSearrch are:  {'var_smoothing': 1.519911082952933e-06}
The confusion matrix and classification report in training set are:
[[ 8  1]
 [ 4 29]]
              precision    recall  f1-score   support

           0       0.89      0.67      0.76        12
           1       0.88      0.97      0.92        30

    accuracy                           0.88        42
   macro avg       0.88      0.82      0.84        42
weighted avg       0.88      0.88      0.88        42

################################################################################################
########################
######                                DECISION TREE MODEL (5ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
DecisionTreeClassifier(class_weight={0: 1, 1: 10}, criterion='entropy',
                       max_depth=100, max_features=1) {'ccp_alpha': 0.0, 'class_weight': {0: 1,
1: 10}, 'criterion': 'entropy', 'max_depth': 100, 'max_features': 1}
The accuracy in training set is: 100.00%
The precision in training set is: 100.00%
```

```
The recall in training set is: 100.00%
The F1 score in training set is: 100.00%
The confusion matrix and classification report in training set are:
[[ 36   0]
 [  0 117]]
              precision    recall  f1-score   support

     Healthy       1.00      1.00      1.00        36
   Parkinson       1.00      1.00      1.00       117

    accuracy                           1.00       153
   macro avg       1.00      1.00      1.00       153
weighted avg       1.00      1.00      1.00       153

The accuracy in testing set is: 83.33%
The precison in testing set is: 87.10%
The recall in testing set is: 90.00%
The F1 score in testing set is: 88.52%
The parameters after the GridSearrch are:  {'ccp_alpha': 0.0, 'class_weight': {0: 1, 1: 10}, 'c
riterion': 'entropy', 'max_depth': 100, 'max_features': 1}
The confusion matrix and classification report in training set are:
[[ 8  3]
 [ 4 27]]
              precision    recall  f1-score   support

           0       0.73      0.67      0.70        12
           1       0.87      0.90      0.89        30

    accuracy                           0.83        42
   macro avg       0.80      0.78      0.79        42
weighted avg       0.83      0.83      0.83        42

##########################################################################################
#######################
######                                LOGISTIC REGRESSION MODEL (5ος διαχωρισμος)
######
######
######
After grid search the Logistic model has the follow parameters:
LogisticRegression(C=1, class_weight={0: 1, 1: 1}, penalty='l1',
                   solver='liblinear', tol=0.01) {'C': 1, 'class_weight': {0: 1, 1: 1}, 'max_it
er': 100, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The accuracy in training set is: 85.62%
The precion in training set is: 86.82%
The recall in training set is: 95.73%
The F1 score in training set is: 91.06%
The confusion matrix and classification report in training set are:
[[ 19   5]
 [ 17 112]]
```

```
              precision    recall  f1-score   support

     Healthy       0.79      0.53      0.63        36
   Parkinson       0.87      0.96      0.91       117

    accuracy                           0.86       153
   macro avg       0.83      0.74      0.77       153
weighted avg       0.85      0.86      0.85       153

The accuracy in testing set is: 85.71%
The precion testing set is: 85.29%
The recall in testing set is: 96.67%
The F1 score in testing set is: 90.62%
The parameters after the GridSearrch are:  {'C': 1, 'class_weight': {0: 1, 1: 1}, 'max_iter': 1
00, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The confusion matrix and classification report in training set are:
[[ 7  1]
 [ 5 29]]
              precision    recall  f1-score   support

           0       0.88      0.58      0.70        12
           1       0.85      0.97      0.91        30

    accuracy                           0.86        42
   macro avg       0.86      0.78      0.80        42
weighted avg       0.86      0.86      0.85        42

##########################################################################################
######################
######                            SUPPORT VECTOR MACHINE MODEL (5ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=0.517947467923121, class_weight='balanced', gamma=1e-09) {'C': 0.517947467923121, 'class_
weight': 'balanced', 'gamma': 1e-09, 'kernel': 'rbf'}
The accuracy in training set is: 76.47%
The precision in training set is: 76.47%
The recall in training set is: 100.00%
The F1 score in training set is: 86.67%
The confusion matrix and classification report in training set are:
[[  0   0]
 [ 36 117]]
              precision    recall  f1-score   support

     Healthy       0.00      0.00      0.00        36
   Parkinson       0.76      1.00      0.87       117

    accuracy                           0.76       153
```

```
   macro avg       0.38      0.50      0.43       153
weighted avg       0.58      0.76      0.66       153

The accuracy in testing set is: 71.43%
The precision in testing set is: 71.43%
The recall in testing set is: 100.00%
The F1 score in testing set is: 83.33%
The parameters after the GridSearrch are:  {'C': 0.517947467923121, 'class_weight': 'balanced',
'gamma': 1e-09, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[ 0  0]
 [12 30]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        12
           1       0.71      1.00      0.83        30

    accuracy                           0.71        42
   macro avg       0.36      0.50      0.42        42
weighted avg       0.51      0.71      0.60        42

[06:42:43] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.

  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core.  Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.


[06:42:43] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
######                                          XGB MODEL (5ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.01, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
              monotone_constraints='()', n_estimators=150, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.3, tree_method='exact',
```

```
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.01, 'n_estimators': 150, 'subsample': 0.3}
The accuracy in training set is: 88.89%
The precision in training set is: 89.68%
The recall in training set is: 96.58%
The F1 score in training set is: 93.00%
The confusion matrix and classification report in training set are:
[[ 23   4]
 [ 13 113]]
              precision    recall  f1-score   support

     Healthy       0.85      0.64      0.73        36
   Parkinson       0.90      0.97      0.93       117

    accuracy                           0.89       153
   macro avg       0.87      0.80      0.83       153
weighted avg       0.89      0.89      0.88       153

The accuracy in testing set is: 88.10%
The precision in testing set is: 87.88%
The recall in testing set is: 96.67%
The F1 score in testing set is: 92.06%
The parameters after the GridSearrch are:  {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.0
1, 'n_estimators': 150, 'subsample': 0.3}
The confusion matrix and classification report in training set are:
[[ 8  1]
 [ 4 29]]
              precision    recall  f1-score   support

           0       0.89      0.67      0.76        12
           1       0.88      0.97      0.92        30

    accuracy                           0.88        42
   macro avg       0.88      0.82      0.84        42
weighted avg       0.88      0.88      0.88        42

##########################################################################################
########################
######                                        RANDOM FOREST MODEL (5ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(class_weight={0: 1, 1: 10}, max_depth=5,
                       min_samples_split=5, n_estimators=300) {'bootstrap': True, 'class_weight
': {0: 1, 1: 10}, 'criterion': 'gini', 'max_depth': 5, 'max_features': 'auto', 'min_samples_spl
it': 5, 'n_estimators': 300, 'oob_score': False}
The accuracy in training set is: 88.24%
The precision in training set is: 86.67%
```

```
The recall in training set is: 100.00%
The F1 score in training set is: 92.86%
The confusion matrix and classification report in training set are:
[[ 18   0]
 [ 18 117]]
              precision    recall  f1-score   support

     Healthy       1.00      0.50      0.67        36
   Parkinson       0.87      1.00      0.93       117

    accuracy                           0.88       153
   macro avg       0.93      0.75      0.80       153
weighted avg       0.90      0.88      0.87       153

The accuracy in testing set is: 85.71%
The precision in testing set is: 83.33%
The recall in testing set is: 100.00%
The F1 score in testing set is: 90.91%
The parameters after the GridSearrch are:  {'bootstrap': True, 'class_weight': {0: 1, 1: 10}, '
criterion': 'gini', 'max_depth': 5, 'max_features': 'auto', 'min_samples_split': 5, 'n_estimato
rs': 300, 'oob_score': False}
The confusion matrix and classification report in training set are:
[[ 6  0]
 [ 6 30]]
              precision    recall  f1-score   support

           0       1.00      0.50      0.67        12
           1       0.83      1.00      0.91        30

    accuracy                           0.86        42
   macro avg       0.92      0.75      0.79        42
weighted avg       0.88      0.86      0.84        42

Στον 6ο διαχωρισμό του dataset έχουμε:
##########################################################################################
######################
######                                GAUSSIAN_NB MODEL (6ος διαχωρισμος)
######
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=0.1873817422860384) {'var_smoothing': 0.1873817422860384}
The accuracy in training set is: 84.97%
The precision in training set is: 85.07%
The recall in training set is: 97.44%
The F1 score in training set is: 90.84%
The confusion matrix and classification report in training set are:
[[ 16   3]
 [ 20 114]]
```

```
              precision    recall  f1-score   support

     Healthy       0.84      0.44      0.58        36
   Parkinson       0.85      0.97      0.91       117

    accuracy                           0.85       153
   macro avg       0.85      0.71      0.75       153
weighted avg       0.85      0.85      0.83       153

The accuracy in testing set is: 78.57%
The precision in testing set is: 76.92%
The recall in testing set is: 100.00%
The F1 score in testing set is: 86.96%
The parameters after the GridSearrch are:  {'var_smoothing': 0.1873817422860384}
The confusion matrix and classification report in training set are:
[[ 3  0]
 [ 9 30]]
              precision    recall  f1-score   support

           0       1.00      0.25      0.40        12
           1       0.77      1.00      0.87        30

    accuracy                           0.79        42
   macro avg       0.88      0.62      0.63        42
weighted avg       0.84      0.79      0.74        42

##########################################################################################
######################
######                                DECISION TREE MODEL (6ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
DecisionTreeClassifier(class_weight={0: 1, 1: 10}, max_depth=2, max_features=1) {'ccp_alpha': 0
.0, 'class_weight': {0: 1, 1: 10}, 'criterion': 'gini', 'max_depth': 2, 'max_features': 1}
The accuracy in training set is: 79.08%
The precision in training set is: 78.52%
The recall in training set is: 100.00%
The F1 score in training set is: 87.97%
The confusion matrix and classification report in training set are:
[[  4   0]
 [ 32 117]]
              precision    recall  f1-score   support

     Healthy       1.00      0.11      0.20        36
   Parkinson       0.79      1.00      0.88       117

    accuracy                           0.79       153
   macro avg       0.89      0.56      0.54       153
```

```
weighted avg       0.84      0.79      0.72       153

The accuracy in testing set is: 71.43%
The precison in testing set is: 71.43%
The recall in testing set is: 100.00%
The F1 score in testing set is: 83.33%
The parameters after the GridSearrch are:  {'ccp_alpha': 0.0, 'class_weight': {0: 1, 1: 10}, 'c
riterion': 'gini', 'max_depth': 2, 'max_features': 1}
The confusion matrix and classification report in training set are:
[[ 0  0]
 [12 30]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        12
           1       0.71      1.00      0.83        30

    accuracy                           0.71        42
   macro avg       0.36      0.50      0.42        42
weighted avg       0.51      0.71      0.60        42

##########################################################################################
######################
######                                LOGISTIC REGRESSION MODEL (6ος διαχωρισμος)
######
######
######
After grid search the Logistic model has the follow parameters:
LogisticRegression(C=100, class_weight={0: 1, 1: 3}, penalty='l1',
                   solver='liblinear', tol=0.01) {'C': 100, 'class_weight': {0: 1, 1: 3}, 'max_
iter': 100, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The accuracy in training set is: 86.93%
The precion in training set is: 85.40%
The recall in training set is: 100.00%
The F1 score in training set is: 92.13%
The confusion matrix and classification report in training set are:
[[ 16   0]
 [ 20 117]]
              precision    recall  f1-score   support

     Healthy       1.00      0.44      0.62        36
   Parkinson       0.85      1.00      0.92       117

    accuracy                           0.87       153
   macro avg       0.93      0.72      0.77       153
weighted avg       0.89      0.87      0.85       153

The accuracy in testing set is: 83.33%
The precion testing set is: 81.08%
The recall in testing set is: 100.00%
```

```
The F1 score in testing set is: 89.55%
The parameters after the GridSearrch are:  {'C': 100, 'class_weight': {0: 1, 1: 3}, 'max_iter':
100, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The confusion matrix and classification report in training set are:
[[ 5  0]
 [ 7 30]]
              precision    recall  f1-score   support

           0       1.00      0.42      0.59        12
           1       0.81      1.00      0.90        30

    accuracy                           0.83        42
   macro avg       0.91      0.71      0.74        42
weighted avg       0.86      0.83      0.81        42

##########################################################################################
######################
######                              SUPPORT VECTOR MACHINE MODEL (6ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=0.517947467923121, class_weight='balanced', gamma=1e-09) {'C': 0.517947467923121, 'class_
weight': 'balanced', 'gamma': 1e-09, 'kernel': 'rbf'}
The accuracy in training set is: 76.47%
The precision in training set is: 76.47%
The recall in training set is: 100.00%
The F1 score in training set is: 86.67%
The confusion matrix and classification report in training set are:
[[  0   0]
 [ 36 117]]
              precision    recall  f1-score   support

     Healthy       0.00      0.00      0.00        36
   Parkinson       0.76      1.00      0.87       117

    accuracy                           0.76       153
   macro avg       0.38      0.50      0.43       153
weighted avg       0.58      0.76      0.66       153

The accuracy in testing set is: 71.43%
The precision in testing set is: 71.43%
The recall in testing set is: 100.00%
The F1 score in testing set is: 83.33%
The parameters after the GridSearrch are:  {'C': 0.517947467923121, 'class_weight': 'balanced',
'gamma': 1e-09, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[ 0  0]
 [12 30]]
```

```
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        12
           1       0.71      1.00      0.83        30

    accuracy                           0.71        42
   macro avg       0.36      0.50      0.42        42
weighted avg       0.51      0.71      0.60        42

[07:15:46] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.

  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core.  Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.


[07:15:46] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
######                                          XGB MODEL (6ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.1, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
              monotone_constraints='()', n_estimators=5, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.5, tree_method='exact',
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.1, 'n_estimators': 5, 'subsample': 0.5}
The accuracy in training set is: 94.12%
The precision in training set is: 93.55%
The recall in training set is: 99.15%
The F1 score in training set is: 96.27%
The confusion matrix and classification report in training set are:
[[ 28   1]
 [  8 116]]
              precision    recall  f1-score   support

     Healthy       0.97      0.78      0.86        36
```

```
   Parkinson       0.94      0.99      0.96       117

    accuracy                           0.94       153
   macro avg       0.95      0.88      0.91       153
weighted avg       0.94      0.94      0.94       153

The accuracy in testing set is: 76.19%
The precision in testing set is: 83.33%
The recall in testing set is: 83.33%
The F1 score in testing set is: 83.33%
The parameters after the GridSearrch are:  {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.1
, 'n_estimators': 5, 'subsample': 0.5}
The confusion matrix and classification report in training set are:
[[ 7  5]
 [ 5 25]]
              precision    recall  f1-score   support

           0       0.58      0.58      0.58        12
           1       0.83      0.83      0.83        30

    accuracy                           0.76        42
   macro avg       0.71      0.71      0.71        42
weighted avg       0.76      0.76      0.76        42

################################################################################################
########################
######                                    RANDOM FOREST MODEL (6ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(class_weight={0: 1, 1: 10}, max_depth=2,
                       min_samples_split=5, n_estimators=500, oob_score=True) {'bootstrap': Tru
e, 'class_weight': {0: 1, 1: 10}, 'criterion': 'gini', 'max_depth': 2, 'max_features': 'auto',
'min_samples_split': 5, 'n_estimators': 500, 'oob_score': True}
The accuracy in training set is: 87.58%
The precision in training set is: 86.03%
The recall in training set is: 100.00%
The F1 score in training set is: 92.49%
The confusion matrix and classification report in training set are:
[[ 17   0]
 [ 19 117]]
              precision    recall  f1-score   support

     Healthy       1.00      0.47      0.64        36
   Parkinson       0.86      1.00      0.92       117

    accuracy                           0.88       153
   macro avg       0.93      0.74      0.78       153
```

```
weighted avg       0.89      0.88      0.86       153

The accuracy in testing set is: 80.95%
The precision in testing set is: 78.95%
The recall in testing set is: 100.00%
The F1 score in testing set is: 88.24%
The parameters after the GridSearrch are:  {'bootstrap': True, 'class_weight': {0: 1, 1: 10}, '
criterion': 'gini', 'max_depth': 2, 'max_features': 'auto', 'min_samples_split': 5, 'n_estimato
rs': 500, 'oob_score': True}
The confusion matrix and classification report in training set are:
[[ 4  0]
 [ 8 30]]
              precision    recall  f1-score   support

           0       1.00      0.33      0.50        12
           1       0.79      1.00      0.88        30

    accuracy                           0.81        42
   macro avg       0.89      0.67      0.69        42
weighted avg       0.85      0.81      0.77        42

Στον 7ο διαχωρισμό του dataset έχουμε:
##########################################################################################
######################
######                                GAUSSIAN_NB MODEL (7ος διαχωρισμος)
######
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=8.111308307896872e-05) {'var_smoothing': 8.111308307896872e-05}
The accuracy in training set is: 81.46%
The precision in training set is: 82.05%
The recall in training set is: 93.20%
The F1 score in training set is: 87.27%
The confusion matrix and classification report in training set are:
[[27  7]
 [21 96]]
              precision    recall  f1-score   support

     Healthy       0.79      0.56      0.66        48
   Parkinson       0.82      0.93      0.87       103

    accuracy                           0.81       151
   macro avg       0.81      0.75      0.77       151
weighted avg       0.81      0.81      0.80       151

The accuracy in testing set is: 77.27%
The precision in testing set is: 100.00%
The recall in testing set is: 77.27%
```

```
The F1 score in testing set is: 87.18%
The parameters after the GridSearrch are:  {'var_smoothing': 8.111308307896872e-05}
The confusion matrix and classification report in training set are:
[[ 0 10]
 [ 0 34]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         0
           1       1.00      0.77      0.87        44

    accuracy                           0.77        44
   macro avg       0.50      0.39      0.44        44
weighted avg       1.00      0.77      0.87        44

##########################################################################################
########################
######                                DECISION TREE MODEL (7ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
DecisionTreeClassifier(class_weight={0: 1, 1: 1}, max_depth=30, max_features=1) {'ccp_alpha': 0
.0, 'class_weight': {0: 1, 1: 1}, 'criterion': 'gini', 'max_depth': 30, 'max_features': 1}
The accuracy in training set is: 100.00%
The precision in training set is: 100.00%
The recall in training set is: 100.00%
The F1 score in training set is: 100.00%
The confusion matrix and classification report in training set are:
[[ 48   0]
 [  0 103]]
              precision    recall  f1-score   support

     Healthy       1.00      1.00      1.00        48
   Parkinson       1.00      1.00      1.00       103

    accuracy                           1.00       151
   macro avg       1.00      1.00      1.00       151
weighted avg       1.00      1.00      1.00       151

The accuracy in testing set is: 88.64%
The precison in testing set is: 100.00%
The recall in testing set is: 88.64%
The F1 score in testing set is: 93.98%
The parameters after the GridSearrch are:  {'ccp_alpha': 0.0, 'class_weight': {0: 1, 1: 1}, 'cr
iterion': 'gini', 'max_depth': 30, 'max_features': 1}
The confusion matrix and classification report in training set are:
[[ 0  5]
 [ 0 39]]
              precision    recall  f1-score   support
```

```
           0       0.00      0.00      0.00         0
           1       1.00      0.89      0.94        44

    accuracy                           0.89        44
   macro avg       0.50      0.44      0.47        44
weighted avg       1.00      0.89      0.94        44

##########################################################################################
######################
######                                LOGISTIC REGRESSION MODEL (7ος διαχωρισμος)
######
######
######
After grid search the Logistic model has the follow parameters:
LogisticRegression(C=100, class_weight={0: 1, 1: 3}, penalty='l1',
                   solver='liblinear', tol=0.01) {'C': 100, 'class_weight': {0: 1, 1: 3}, 'max_
iter': 100, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The accuracy in training set is: 84.11%
The precion in training set is: 81.10%
The recall in training set is: 100.00%
The F1 score in training set is: 89.57%
The confusion matrix and classification report in training set are:
[[ 24   0]
 [ 24 103]]
              precision    recall  f1-score   support

     Healthy       1.00      0.50      0.67        48
   Parkinson       0.81      1.00      0.90       103

    accuracy                           0.84       151
   macro avg       0.91      0.75      0.78       151
weighted avg       0.87      0.84      0.82       151

The accuracy in testing set is: 97.73%
The precion testing set is: 100.00%
The recall in testing set is: 97.73%
The F1 score in testing set is: 98.85%
The parameters after the GridSearrch are:  {'C': 100, 'class_weight': {0: 1, 1: 3}, 'max_iter':
100, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The confusion matrix and classification report in training set are:
[[ 0  1]
 [ 0 43]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         0
           1       1.00      0.98      0.99        44

    accuracy                           0.98        44
```

```
   macro avg       0.50      0.49      0.49        44
weighted avg       1.00      0.98      0.99        44

##########################################################################################
######################
######                              SUPPORT VECTOR MACHINE MODEL (7ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=10000000000.0, class_weight={0: 1, 1: 3}, gamma=2.1544346900318822e-08) {'C': 10000000000
.0, 'class_weight': {0: 1, 1: 3}, 'gamma': 2.1544346900318822e-08, 'kernel': 'rbf'}
The accuracy in training set is: 83.44%
The precision in training set is: 81.97%
The recall in training set is: 97.09%
The F1 score in training set is: 88.89%
The confusion matrix and classification report in training set are:
[[ 26   3]
 [ 22 100]]
              precision    recall  f1-score   support

     Healthy       0.90      0.54      0.68        48
   Parkinson       0.82      0.97      0.89       103

    accuracy                           0.83       151
   macro avg       0.86      0.76      0.78       151
weighted avg       0.84      0.83      0.82       151

The accuracy in testing set is: 95.45%
The precision in testing set is: 100.00%
The recall in testing set is: 95.45%
The F1 score in testing set is: 97.67%
The parameters after the GridSearrch are:  {'C': 10000000000.0, 'class_weight': {0: 1, 1: 3}, '
gamma': 2.1544346900318822e-08, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[ 0  2]
 [ 0 42]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         0
           1       1.00      0.95      0.98        44

    accuracy                           0.95        44
   macro avg       0.50      0.48      0.49        44
weighted avg       1.00      0.95      0.98        44

[07:49:56] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.
```

```
  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core.  Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.


[07:49:56] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
######                                          XGB MODEL (7ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.5, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
              monotone_constraints='()', n_estimators=200, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.3, tree_method='exact',
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.5, 'n_estimators': 200, 'subsample': 0.3}
The accuracy in training set is: 96.69%
The precision in training set is: 98.04%
The recall in training set is: 97.09%
The F1 score in training set is: 97.56%
The confusion matrix and classification report in training set are:
[[ 46   3]
 [  2 100]]
              precision    recall  f1-score   support

     Healthy       0.94      0.96      0.95        48
   Parkinson       0.98      0.97      0.98       103

    accuracy                           0.97       151
   macro avg       0.96      0.96      0.96       151
weighted avg       0.97      0.97      0.97       151

The accuracy in testing set is: 84.09%
The precision in testing set is: 100.00%
The recall in testing set is: 84.09%
The F1 score in testing set is: 91.36%
The parameters after the GridSearrch are:  {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.5
, 'n_estimators': 200, 'subsample': 0.3}
```

```
The confusion matrix and classification report in training set are:
[[ 0  7]
 [ 0 37]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         0
           1       1.00      0.84      0.91        44

    accuracy                           0.84        44
   macro avg       0.50      0.42      0.46        44
weighted avg       1.00      0.84      0.91        44

##########################################################################################
######################
######                                        RANDOM FOREST MODEL (7ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(bootstrap=False, class_weight={0: 1, 1: 3},
                       criterion='entropy', max_depth=5, n_estimators=500) {'bootstrap': False,
'class_weight': {0: 1, 1: 3}, 'criterion': 'entropy', 'max_depth': 5, 'max_features': 'auto', '
min_samples_split': 2, 'n_estimators': 500, 'oob_score': False}
The accuracy in training set is: 97.35%
The precision in training set is: 96.26%
The recall in training set is: 100.00%
The F1 score in training set is: 98.10%
The confusion matrix and classification report in training set are:
[[ 44   0]
 [  4 103]]
              precision    recall  f1-score   support

     Healthy       1.00      0.92      0.96        48
   Parkinson       0.96      1.00      0.98       103

    accuracy                           0.97       151
   macro avg       0.98      0.96      0.97       151
weighted avg       0.97      0.97      0.97       151

The accuracy in testing set is: 100.00%
The precision in testing set is: 100.00%
The recall in testing set is: 100.00%
The F1 score in testing set is: 100.00%
The parameters after the GridSearrch are:  {'bootstrap': False, 'class_weight': {0: 1, 1: 3}, '
criterion': 'entropy', 'max_depth': 5, 'max_features': 'auto', 'min_samples_split': 2, 'n_estim
ators': 500, 'oob_score': False}
The confusion matrix and classification report in training set are:
[[44]]
              precision    recall  f1-score   support
```

```
           1       1.00      1.00      1.00        44

    accuracy                           1.00        44
   macro avg       1.00      1.00      1.00        44
weighted avg       1.00      1.00      1.00        44

Στον 8ο διαχωρισμό του dataset έχουμε:
################################################################################################
######################
######                                  GAUSSIAN_NB MODEL (8ος διαχωρισμος)
######
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=0.3511191734215131) {'var_smoothing': 0.3511191734215131}
The accuracy in training set is: 84.87%
The precision in training set is: 83.45%
The recall in training set is: 100.00%
The F1 score in training set is: 90.98%
The confusion matrix and classification report in training set are:
[[ 13   0]
 [ 23 116]]
              precision    recall  f1-score   support

     Healthy       1.00      0.36      0.53        36
   Parkinson       0.83      1.00      0.91       116

    accuracy                           0.85       152
   macro avg       0.92      0.68      0.72       152
weighted avg       0.87      0.85      0.82       152

The accuracy in testing set is: 72.09%
The precision in testing set is: 72.09%
The recall in testing set is: 100.00%
The F1 score in testing set is: 83.78%
The parameters after the GridSearrch are:  {'var_smoothing': 0.3511191734215131}
The confusion matrix and classification report in training set are:
[[ 0  0]
 [12 31]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        12
           1       0.72      1.00      0.84        31

    accuracy                           0.72        43
   macro avg       0.36      0.50      0.42        43
weighted avg       0.52      0.72      0.60        43
```

```
##########################################################################################
######################
######                                DECISION TREE MODEL (8ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
DecisionTreeClassifier(class_weight={0: 1, 1: 1}, criterion='entropy',
                       max_depth=10, max_features=1) {'ccp_alpha': 0.0, 'class_weight': {0: 1,
1: 1}, 'criterion': 'entropy', 'max_depth': 10, 'max_features': 1}
The accuracy in training set is: 99.34%
The precision in training set is: 100.00%
The recall in training set is: 99.14%
The F1 score in training set is: 99.57%
The confusion matrix and classification report in training set are:
[[ 36   1]
 [  0 115]]
              precision    recall  f1-score   support

     Healthy       0.97      1.00      0.99        36
   Parkinson       1.00      0.99      1.00       116

    accuracy                           0.99       152
   macro avg       0.99      1.00      0.99       152
weighted avg       0.99      0.99      0.99       152

The accuracy in testing set is: 86.05%
The precison in testing set is: 83.78%
The recall in testing set is: 100.00%
The F1 score in testing set is: 91.18%
The parameters after the GridSearrch are:  {'ccp_alpha': 0.0, 'class_weight': {0: 1, 1: 1}, 'cr
iterion': 'entropy', 'max_depth': 10, 'max_features': 1}
The confusion matrix and classification report in training set are:
[[ 6  0]
 [ 6 31]]
              precision    recall  f1-score   support

           0       1.00      0.50      0.67        12
           1       0.84      1.00      0.91        31

    accuracy                           0.86        43
   macro avg       0.92      0.75      0.79        43
weighted avg       0.88      0.86      0.84        43

##########################################################################################
######################
######                                LOGISTIC REGRESSION MODEL (8ος διαχωρισμος)
######
```

```
######
######
After grid search the Logistic model has the follow parameters:
LogisticRegression(C=100, class_weight={0: 1, 1: 3}, penalty='l1',
                   solver='liblinear', tol=0.01) {'C': 100, 'class_weight': {0: 1, 1: 3}, 'max_
iter': 100, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The accuracy in training set is: 88.82%
The precion in training set is: 87.22%
The recall in training set is: 100.00%
The F1 score in training set is: 93.17%
The confusion matrix and classification report in training set are:
[[ 19   0]
 [ 17 116]]
              precision    recall  f1-score   support

     Healthy       1.00      0.53      0.69        36
   Parkinson       0.87      1.00      0.93       116

    accuracy                           0.89       152
   macro avg       0.94      0.76      0.81       152
weighted avg       0.90      0.89      0.87       152

The accuracy in testing set is: 83.72%
The precion testing set is: 81.58%
The recall in testing set is: 100.00%
The F1 score in testing set is: 89.86%
The parameters after the GridSearrch are:  {'C': 100, 'class_weight': {0: 1, 1: 3}, 'max_iter':
100, 'penalty': 'l1', 'solver': 'liblinear', 'tol': 0.01}
The confusion matrix and classification report in training set are:
[[ 5  0]
 [ 7 31]]
              precision    recall  f1-score   support

           0       1.00      0.42      0.59        12
           1       0.82      1.00      0.90        31

    accuracy                           0.84        43
   macro avg       0.91      0.71      0.74        43
weighted avg       0.87      0.84      0.81        43

##############################################################################################
######################
######                              SUPPORT VECTOR MACHINE MODEL (8ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=26.826957952797247, class_weight='balanced', gamma=1e-09) {'C': 26.826957952797247, 'clas
s_weight': 'balanced', 'gamma': 1e-09, 'kernel': 'rbf'}
```

```
The accuracy in training set is: 23.68%
The precision in training set is: 0.00%
The recall in training set is: 0.00%
The F1 score in training set is: 0.00%
The confusion matrix and classification report in training set are:
[[ 36 116]
 [  0   0]]
              precision    recall  f1-score   support

     Healthy       0.24      1.00      0.38        36
   Parkinson       0.00      0.00      0.00       116

    accuracy                           0.24       152
   macro avg       0.12      0.50      0.19       152
weighted avg       0.06      0.24      0.09       152

The accuracy in testing set is: 27.91%
The precision in testing set is: 0.00%
The recall in testing set is: 0.00%
The F1 score in testing set is: 0.00%
The parameters after the GridSearrch are:  {'C': 26.826957952797247, 'class_weight': 'balanced'
, 'gamma': 1e-09, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[12 31]
 [ 0  0]]
              precision    recall  f1-score   support

           0       0.28      1.00      0.44        12
           1       0.00      0.00      0.00        31

    accuracy                           0.28        43
   macro avg       0.14      0.50      0.22        43
weighted avg       0.08      0.28      0.12        43

[08:22:46] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.

  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core.  Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.


[08:22:46] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
```

XGB MODEL (8ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.5, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
              monotone_constraints='()', n_estimators=150, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.5, tree_method='exact',
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.5, 'n_estimators': 150, 'subsample': 0.5}
The accuracy in training set is: 100.00%
The precision in training set is: 100.00%
The recall in training set is: 100.00%
The F1 score in training set is: 100.00%
The confusion matrix and classification report in training set are:
[[36   0]
 [0 116]]
              precision   recall f1-score   support

     Healthy   1.00   1.00   1.00   36
   Parkinson   1.00   1.00   1.00   116

    accuracy   1.00   152
   macro avg   1.00   1.00   1.00   152
weighted avg   1.00   1.00   1.00   152

The accuracy in testing set is: 88.37%
The precision in testing set is: 90.62%
The recall in testing set is: 93.55%
The F1 score in testing set is: 92.06%
The parameters after the GridSearrch are: {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.5
, 'n_estimators': 150, 'subsample': 0.5}
The confusion matrix and classification report in training set are:
[[9 2]
 [3 29]]
              precision   recall f1-score   support

           0   0.82   0.75   0.78   12
           1   0.91   0.94   0.92   31

    accuracy   0.88   43
   macro avg   0.86   0.84   0.85   43
weighted avg   0.88   0.88   0.88   43

```
##########################################################################################
######################
######                                    RANDOM FOREST MODEL (8ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(class_weight={0: 1, 1: 1}, max_depth=2, n_estimators=500) {'bootstrap':
True, 'class_weight': {0: 1, 1: 1}, 'criterion': 'gini', 'max_depth': 2, 'max_features': 'auto'
, 'min_samples_split': 2, 'n_estimators': 500, 'oob_score': False}
The accuracy in training set is: 89.47%
The precision in training set is: 87.88%
The recall in training set is: 100.00%
The F1 score in training set is: 93.55%
The confusion matrix and classification report in training set are:
[[ 20   0]
 [ 16 116]]
              precision    recall  f1-score   support

     Healthy       1.00      0.56      0.71        36
   Parkinson       0.88      1.00      0.94       116

    accuracy                           0.89       152
   macro avg       0.94      0.78      0.82       152
weighted avg       0.91      0.89      0.88       152

The accuracy in testing set is: 72.09%
The precision in testing set is: 72.09%
The recall in testing set is: 100.00%
The F1 score in testing set is: 83.78%
The parameters after the GridSearrch are:  {'bootstrap': True, 'class_weight': {0: 1, 1: 1}, 'c
riterion': 'gini', 'max_depth': 2, 'max_features': 'auto', 'min_samples_split': 2, 'n_estimator
s': 500, 'oob_score': False}
The confusion matrix and classification report in training set are:
[[ 0  0]
 [12 31]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        12
           1       0.72      1.00      0.84        31

    accuracy                           0.72        43
   macro avg       0.36      0.50      0.42        43
weighted avg       0.52      0.72      0.60        43

Στον 9ο διαχωρισμό του dataset έχουμε:
##########################################################################################
######################
```

```
######                                GAUSSIAN_NB MODEL (9ος διαχωρισμος)
######
######
######
After grid search the Naive Bayes model has the follow parameters:
GaussianNB(var_smoothing=4.328761281083062e-05) {'var_smoothing': 4.328761281083062e-05}
The accuracy in training set is: 82.35%
The precision in training set is: 81.45%
The recall in training set is: 96.19%
The F1 score in training set is: 88.21%
The confusion matrix and classification report in training set are:
[[ 25   4]
 [ 23 101]]
              precision    recall  f1-score   support

     Healthy       0.86      0.52      0.65        48
   Parkinson       0.81      0.96      0.88       105

    accuracy                           0.82       153
   macro avg       0.84      0.74      0.77       153
weighted avg       0.83      0.82      0.81       153

The accuracy in testing set is: 88.10%
The precision in testing set is: 100.00%
The recall in testing set is: 88.10%
The F1 score in testing set is: 93.67%
The parameters after the GridSearrch are:  {'var_smoothing': 4.328761281083062e-05}
The confusion matrix and classification report in training set are:
[[ 0  5]
 [ 0 37]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         0
           1       1.00      0.88      0.94        42

    accuracy                           0.88        42
   macro avg       0.50      0.44      0.47        42
weighted avg       1.00      0.88      0.94        42

#########################################################################################
########################
######                                DECISION TREE MODEL (9ος διαχωρισμος)
######
######
######
After grid search the Decision Tree model has the follow parameters:
DecisionTreeClassifier(class_weight={0: 1, 1: 1}, max_depth=100, max_features=1) {'ccp_alpha':
0.0, 'class_weight': {0: 1, 1: 1}, 'criterion': 'gini', 'max_depth': 100, 'max_features': 1}
The accuracy in training set is: 100.00%
```

```
The precision in training set is: 100.00%
The recall in training set is: 100.00%
The F1 score in training set is: 100.00%
The confusion matrix and classification report in training set are:
[[ 48   0]
 [  0 105]]
              precision    recall  f1-score   support

     Healthy       1.00      1.00      1.00        48
   Parkinson       1.00      1.00      1.00       105

    accuracy                           1.00       153
   macro avg       1.00      1.00      1.00       153
weighted avg       1.00      1.00      1.00       153

The accuracy in testing set is: 85.71%
The precison in testing set is: 100.00%
The recall in testing set is: 85.71%
The F1 score in testing set is: 92.31%
The parameters after the GridSearrch are:  {'ccp_alpha': 0.0, 'class_weight': {0: 1, 1: 1}, 'cr
iterion': 'gini', 'max_depth': 100, 'max_features': 1}
The confusion matrix and classification report in training set are:
[[ 0  6]
 [ 0 36]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         0
           1       1.00      0.86      0.92        42

    accuracy                           0.86        42
   macro avg       0.50      0.43      0.46        42
weighted avg       1.00      0.86      0.92        42

##########################################################################################
######################
######                                LOGISTIC REGRESSION MODEL (9ος διαχωρισμος)
######
######
######
After grid search the Logistic model has the follow parameters:
LogisticRegression(C=1, class_weight={0: 1, 1: 3}, max_iter=1000, tol=1e-08) {'C': 1, 'class_we
ight': {0: 1, 1: 3}, 'max_iter': 1000, 'penalty': 'l2', 'solver': 'lbfgs', 'tol': 1e-08}
The accuracy in training set is: 84.97%
The precion in training set is: 82.54%
The recall in training set is: 99.05%
The F1 score in training set is: 90.04%
The confusion matrix and classification report in training set are:
[[ 26   1]
 [ 22 104]]
```

```
              precision    recall  f1-score   support

     Healthy       0.96      0.54      0.69        48
   Parkinson       0.83      0.99      0.90       105

    accuracy                           0.85       153
   macro avg       0.89      0.77      0.80       153
weighted avg       0.87      0.85      0.84       153

The accuracy in testing set is: 97.62%
The precion testing set is: 100.00%
The recall in testing set is: 97.62%
The F1 score in testing set is: 98.80%
The parameters after the GridSearrch are:  {'C': 1, 'class_weight': {0: 1, 1: 3}, 'max_iter': 1
000, 'penalty': 'l2', 'solver': 'lbfgs', 'tol': 1e-08}
The confusion matrix and classification report in training set are:
[[ 0  1]
 [ 0 41]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         0
           1       1.00      0.98      0.99        42

    accuracy                           0.98        42
   macro avg       0.50      0.49      0.49        42
weighted avg       1.00      0.98      0.99        42

##########################################################################################
#######################
######                            SUPPORT VECTOR MACHINE MODEL (9ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
SVC(C=0.01, class_weight='balanced', gamma=1e-09) {'C': 0.01, 'class_weight': 'balanced', 'gamm
a': 1e-09, 'kernel': 'rbf'}
The accuracy in training set is: 31.37%
The precision in training set is: 0.00%
The recall in training set is: 0.00%
The F1 score in training set is: 0.00%
The confusion matrix and classification report in training set are:
[[ 48 105]
 [  0   0]]
              precision    recall  f1-score   support

     Healthy       0.31      1.00      0.48        48
   Parkinson       0.00      0.00      0.00       105

    accuracy                           0.31       153
```

```
   macro avg       0.16      0.50      0.24       153
weighted avg       0.10      0.31      0.15       153

The accuracy in testing set is: 0.00%
The precision in testing set is: 0.00%
The recall in testing set is: 0.00%
The F1 score in testing set is: 0.00%
The parameters after the GridSearrch are:  {'C': 0.01, 'class_weight': 'balanced', 'gamma': 1e-
09, 'kernel': 'rbf'}
The confusion matrix and classification report in training set are:
[[ 0 42]
 [ 0  0]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       0.0
           1       0.00      0.00      0.00      42.0

    accuracy                           0.00      42.0
   macro avg       0.00      0.00      0.00      42.0
weighted avg       0.00      0.00      0.00      42.0

[08:57:33] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:541:
Parameters: { class_weight } might not be used.

  This may not be accurate due to some parameters are only used in language bindings but
  passed down to XGBoost core.  Or some parameters are not used but slip through this
  verification. Please open an issue if you find above cases.


[08:57:33] WARNING: C:/Users/Administrator/workspace/xgboost-win64_release_1.3.0/src/learner.cc
:1061: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary
:logistic' was changed from 'error' to 'logloss'. Explicitly set eval_metric if you'd like to r
estore the old behavior.
################################################################################################
######################
######                                          XGB MODEL (9ος διαχωρισμος)
######
######
######
After grid search the SVM model has the follow parameters:
XGBClassifier(base_score=0.5, booster='gbtree', class_weight={0: 10, 1: 1},
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              gamma=0, gpu_id=-1, importance_type='gain',
              interaction_constraints='', learning_rate=0.9, max_delta_step=0,
              max_depth=6, min_child_weight=1, missing=nan,
              monotone_constraints='()', n_estimators=200, n_jobs=12,
              num_parallel_tree=1, random_state=0, reg_alpha=0, reg_lambda=1,
              scale_pos_weight=1, subsample=0.3, tree_method='exact',
```

```
              validate_parameters=1, verbosity=None) {'class_weight': {0: 10, 1: 1}, 'learning_
rate': 0.9, 'n_estimators': 200, 'subsample': 0.3}
The accuracy in training set is: 99.35%
The precision in training set is: 100.00%
The recall in training set is: 99.05%
The F1 score in training set is: 99.52%
The confusion matrix and classification report in training set are:
[[ 48   1]
 [  0 104]]
              precision    recall  f1-score   support

     Healthy       0.98      1.00      0.99        48
   Parkinson       1.00      0.99      1.00       105

    accuracy                           0.99       153
   macro avg       0.99      1.00      0.99       153
weighted avg       0.99      0.99      0.99       153

The accuracy in testing set is: 88.10%
The precision in testing set is: 100.00%
The recall in testing set is: 88.10%
The F1 score in testing set is: 93.67%
The parameters after the GridSearrch are:  {'class_weight': {0: 10, 1: 1}, 'learning_rate': 0.9
, 'n_estimators': 200, 'subsample': 0.3}
The confusion matrix and classification report in training set are:
[[ 0  5]
 [ 0 37]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         0
           1       1.00      0.88      0.94        42

    accuracy                           0.88        42
   macro avg       0.50      0.44      0.47        42
weighted avg       1.00      0.88      0.94        42

##########################################################################################
########################
######                                        RANDOM FOREST MODEL (9ος διαχωρισμος)
######
######
######
After grid search the RFC model has the follow parameters:
RandomForestClassifier(bootstrap=False, class_weight='balanced', max_depth=5) {'bootstrap': Fal
se, 'class_weight': 'balanced', 'criterion': 'gini', 'max_depth': 5, 'max_features': 'auto', 'm
in_samples_split': 2, 'n_estimators': 100, 'oob_score': False}
The accuracy in training set is: 100.00%
The precision in training set is: 100.00%
The recall in training set is: 100.00%
```

```
The F1 score in training set is: 100.00%
The confusion matrix and classification report in training set are:
[[ 48   0]
 [  0 105]]
              precision    recall  f1-score   support

     Healthy       1.00      1.00      1.00        48
   Parkinson       1.00      1.00      1.00       105

    accuracy                           1.00       153
   macro avg       1.00      1.00      1.00       153
weighted avg       1.00      1.00      1.00       153

The accuracy in testing set is: 90.48%
The precision in testing set is: 100.00%
The recall in testing set is: 90.48%
The F1 score in testing set is: 95.00%
The parameters after the GridSearrch are:  {'bootstrap': False, 'class_weight': 'balanced', 'cr
iterion': 'gini', 'max_depth': 5, 'max_features': 'auto', 'min_samples_split': 2, 'n_estimators
': 100, 'oob_score': False}
The confusion matrix and classification report in training set are:
[[ 0  4]
 [ 0 38]]
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         0
           1       1.00      0.90      0.95        42

    accuracy                           0.90        42
   macro avg       0.50      0.45      0.48        42
weighted avg       1.00      0.90      0.95        42
```