In [3]:
```
#Data Set 1: adult.data.csv
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.utils import shuffle
from sklearn import preprocessing
from sklearn.preprocessing import LabelEncoder
%config InlineBackend.figure_format = 'retina'

# 1) load file
fp1 = 'adult.data.csv'
df1 = pd.read_csv(fp1)

# 2) preprocess data: one-hot encoding using sklearn package
le = preprocessing.LabelEncoder()
data1 = df1.apply(le.fit_transform)
data1 = data1.values
np.random.shuffle(data1)
print(data1.shape)
# label data and split them
X_data1 = data1[:, 0:14]
Y_data1 = data1[:, 14]

#3) split data into three partitions
#partition 1: 0.8 training, 0.2 testing
partition1 = int(0.8*len(data1))
train11_X = X_data1[:partition1, :]
train11_Y = Y_data1[:partition1]
test11_X = X_data1[partition1:, :]
test11_Y = Y_data1[partition1:]
print("train1_X has shape", train11_X.shape, ", train1_Y has shape", train11_Y.shape)
print("test1_X has shape", test11_X.shape, ", test1_Y has shape", test11_Y.shape)
#partition 2: 0.5 training, 0.5 testing
partition2 = int(0.5*len(data1))
train12_X = X_data1[:partition2, :]
train12_Y = Y_data1[:partition2]
test12_X = X_data1[partition2:, :]
test12_Y = Y_data1[partition2:]
print("train2_X has shape", train12_X.shape, ", train2_Y has shape", train12_Y.shape)
print("test2_X has shape", test12_X.shape, ", test2_Y has shape", test12_Y.shape)
#partition 3: 0.2 training, 0.8 testing
partition3 = int(0.2*len(data1))
train13_X = X_data1[:partition3, :]
train13_Y = Y_data1[:partition3]
test13_X = X_data1[partition3:, :]
test13_Y = Y_data1[partition3:]
print("train3_X has shape", train13_X.shape, ", train3_Y has shape", train13_Y.shape)
print("test3_X has shape", test13_X.shape, ", test3_Y has shape", test13_Y.shape)
```

```
(32561, 15)
train1_X has shape (26048, 14) , train1_Y has shape (26048,)
test1_X has shape (6513, 14) , test1_Y has shape (6513,)
train2_X has shape (16280, 14) , train2_Y has shape (16280,)
test2_X has shape (16281, 14) , test2_Y has shape (16281,)
train3_X has shape (6512, 14) , train3_Y has shape (6512,)
test3_X has shape (26049, 14) , test3_Y has shape (26049,)
```

In [4]:
```
#Data Set 2: car.data.csv
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.utils import shuffle
from sklearn import preprocessing
from sklearn.preprocessing import LabelEncoder

# 1) load file
fp2 = 'car.data.csv'
df2 = pd.read_csv(fp2)

# 2) preprocess data: one-hot encoding using sklearn package
le = preprocessing.LabelEncoder()
data2 = df2.apply(le.fit_transform)
data2 = data2.values
np.random.shuffle(data2)
print(data2.shape)
# label data and split them
X_data = data2[:, 0:6]
Y_data = data2[:, 6]

#3) split data into three partitions
#partition 1: 0.8 training, 0.2 testing
partition1 = int(0.8*len(data2))
train21_X = X_data[:partition1, :]
train21_Y = Y_data[:partition1]
test21_X = X_data[partition1:, :]
test21_Y = Y_data[partition1:]
print("train1_X has shape", train21_X.shape, ", train1_Y has shape", train21_Y.shape)
print("test1_X has shape", test21_X.shape, ", test1_Y has shape", test21_Y.shape)
#partition 2: 0.5 training, 0.5 testing
partition2 = int(0.5*len(data2))
train22_X = X_data[:partition2, :]
train22_Y = Y_data[:partition2]
test22_X = X_data[partition2:, :]
test22_Y = Y_data[partition2:]
print("train2_X has shape", train22_X.shape, ", train2_Y has shape", train22_Y.shape)
print("test2_X has shape", test22_X.shape, ", test2_Y has shape", test22_Y.shape)
#partition 3: 0.2 training, 0.8 testing
partition3 = int(0.2*len(data2))
train23_X = X_data[:partition3, :]
train23_Y = Y_data[:partition3]
test23_X = X_data[partition3:, :]
test23_Y = Y_data[partition3:]
print("train3_X has shape", train23_X.shape, ", train3_Y has shape", train23_Y.shape)
print("test3_X has shape", test23_X.shape, ", test3_Y has shape", test23_Y.shape)
```

```
(1727, 7)
train1_X has shape (1381, 6) , train1_Y has shape (1381,)
test1_X has shape (346, 6) , test1_Y has shape (346,)
train2_X has shape (863, 6) , train2_Y has shape (863,)
test2_X has shape (864, 6) , test2_Y has shape (864,)
train3_X has shape (345, 6) , train3_Y has shape (345,)
test3_X has shape (1382, 6) , test3_Y has shape (1382,)
```

In [5]:
```
#Data Set 3: winequality-red.csv
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.utils import shuffle
from sklearn import preprocessing
from sklearn.preprocessing import LabelEncoder

# 1) load file
fp3 = 'winequality-red.csv'
df3 = pd.read_csv(fp3, sep = ';')

# 2) preprocess data: one-hot encoding using sklearn package
le = preprocessing.LabelEncoder()
data3 = df3.apply(le.fit_transform)
data3 = data3.values
np.random.shuffle(data3)
print(data3.shape)
# label data and split them
X_data = data3[:, 0:11]
Y_data = data3[:, 11]

#3) split data into three partitions
#partition 1: 0.8 training, 0.2 testing
partition1 = int(0.8*len(data3))
train31_X = X_data[:partition1, :]
train31_Y = Y_data[:partition1]
test31_X = X_data[partition1:, :]
test31_Y = Y_data[partition1:]
print("train1_X has shape", train31_X.shape, ", train1_Y has shape", train31_Y.shape)
print("test1_X has shape", test31_X.shape, ", test1_Y has shape", test31_Y.shape)
#partition 2: 0.5 training, 0.5 testing
partition2 = int(0.5*len(data3))
train32_X = X_data[:partition2, :]
train32_Y = Y_data[:partition2]
test32_X = X_data[partition2:, :]
test32_Y = Y_data[partition2:]
print("train2_X has shape", train32_X.shape, ", train2_Y has shape", train32_Y.shape)
print("test2_X has shape", test32_X.shape, ", test2_Y has shape", test32_Y.shape)
#partition 3: 0.2 training, 0.8 testing
partition3 = int(0.2*len(data3))
train33_X = X_data[:partition3, :]
train33_Y = Y_data[:partition3]
test33_X = X_data[partition3:, :]
test33_Y = Y_data[partition3:]
print("train3_X has shape", train33_X.shape, ", train3_Y has shape", train33_Y.shape)
print("test3_X has shape", test33_X.shape, ", test3_Y has shape", test33_Y.shape)
```

```
(1599, 12)
train1_X has shape (1279, 11) , train1_Y has shape (1279,)
test1_X has shape (320, 11) , test1_Y has shape (320,)
train2_X has shape (799, 11) , train2_Y has shape (799,)
test2_X has shape (800, 11) , test2_Y has shape (800,)
train3_X has shape (319, 11) , train3_Y has shape (319,)
test3_X has shape (1280, 11) , test3_Y has shape (1280,)
```

Start training classifiers for three data sets

In [10]:
```
# import Classifier1_NNMLP
# from Classifier1_NNMLP import NN_MLP
# from Classfier2_RF import random_forest_classifier
# from Classfier3_GradientBoosting import gradient_boosting_classifier
import multiprocessing
from sklearn.neural_network import MLPClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import RandomizedSearchCV
from sklearn.model_selection import GridSearchCV
from sklearn.metrics import classification_report
import warnings
warnings.filterwarnings("ignore")
def NN_MLP(X_train, Y_train, X_test, Y_test):
# preprocess data since MLP is sensitive to feature scaling

    scaler = StandardScaler()
    scaler.fit(X_train)
    X_train = scaler.transform(X_train)
    X_test = scaler.transform(X_test)

    test_accs = []

    for i in range(0, 3):
    #parameters
    # solver uses adam here since it works better on large data sets
    # activation use relu since the label are 0 and 1
    # hidden layer size : input_size->3000->1000->100 ->1
    # use gridsearchcv to choose best alpha -> regularzation
        mlp = MLPClassifier(max_iter = 1000, hidden_layer_sizes = (100,100))
        parameters = {'alpha':10.0 ** -np.arange(1,7), 'learning_rate': ["constant",
 "invscaling", "adaptive"], 'solver':['sgd', 'adam', 'lbfgs']}
        clf = GridSearchCV(mlp, parameters, n_jobs = multiprocessing.cpu_count
()-1, cv = 3)
        clf.fit(X_train, Y_train)
        print("best estimator is: ", clf.best_estimator_)
        optimal_mlp = MLPClassifier(solver = clf.best_params_['solver'],max_iter =
 1000, learning_rate = clf.best_params_['learning_rate'], hidden_layer_sizes = (100,
100), alpha =clf.best_params_['alpha'] )
        optimal_mlp.fit(X_train, Y_train)
        predictions = optimal_mlp.predict(X_test)
        test_acc = optimal_mlp.score(X_test, Y_test)
        test_accs.append(test_acc)
        print("test accuracy is:", test_acc)
        print(classification_report(Y_test,predictions))

        train_acc = clf.cv_results_['mean_train_score']
        val_acc = clf.cv_results_['mean_test_score']
        print("The accuarcy for cross_validation")
        plt.plot(train_acc, label = 'train_acc')
        plt.plot(val_acc, label = 'val_acc')
        plt.legend(loc = 'upper left')
        plt.show()
```

```
    print("average of test_accuracy is:", np.mean(test_accs))

def random_forest_classifier(X_train, Y_train, X_test, Y_test):
    #preprocess data
    scaler = StandardScaler()
    scaler.fit(X_train)
    X_train = scaler.transform(X_train)
    X_test = scaler.transform(X_test)

    #train classifier
    n_estimators = [int(x) for x in np.linspace(start = 200, stop = 2000, num
= 10)]
    max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]
    max_depth.append(None)
    max_features = ['auto', 'sqrt']
    parameters = {'n_estimators': n_estimators, 'max_depth': max_depth, 'max_fea
tures': max_features}

    test_accs = []

    for i in range(0, 3):
        rfc = RandomForestClassifier()
        rf = RandomizedSearchCV(rfc, parameters, n_jobs = multiprocessing.cpu_
count()-1, cv = 3)
        rf.fit(X_train, Y_train)
        print("best estimator is: ", rf.best_estimator_)
        optimal_rfc =RandomForestClassifier(n_estimators = rf.best_params_['n_est
imators'], max_depth = rf.best_params_['max_depth'], max_features = rf.best_para
ms_['max_features'])
        optimal_rfc.fit(X_train, Y_train)
        predictions = optimal_rfc.predict(X_test)
        test_acc = optimal_rfc.score(X_test, Y_test)
        test_accs.append(test_acc)
        print("test accuracy is:", test_acc)
        print(classification_report(Y_test,predictions))

        train_acc = rf.cv_results_['mean_train_score']
        val_acc = rf.cv_results_['mean_test_score']
        print("The accuarcy for cross_validation")
        plt.plot(train_acc, label = 'train_acc')
        plt.plot(val_acc, label = 'val_acc')
        plt.legend(loc = 'upper left')
        plt.show()
    print("average of test_accuracy is:", np.mean(test_accs))

def gb_classifier(X_train, Y_train, X_test, Y_test):
    #preprocess data
    scaler = StandardScaler()
    scaler.fit(X_train)
    X_train = scaler.transform(X_train)
    X_test = scaler.transform(X_test)

    test_accs = []
```

```
    for i in range(0, 3):
        n_estimators = [int(x) for x in np.linspace(start = 200, stop = 2000, n
um = 10)]
        max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]
        learning_rate = [0.15,0.1,0.05,0.01,0.005,0.001]
        parameters = {'n_estimators':n_estimators, 'max_depth':max_depth,'learnin
g_rate':learning_rate}

        gbc = GradientBoostingClassifier()
        gb = RandomizedSearchCV(gbc, parameters, n_jobs = multiprocessing.cp
u_count()-1, cv = 3)
        gb.fit(X_train, Y_train)
        print("best estimator is: ", gb.best_estimator_)
        optimal_gbc =GradientBoostingClassifier(n_estimators = gb.best_params_[
'n_estimators'], max_depth = gb.best_params_['max_depth'], learning_rate = gb.bes
t_params_['learning_rate'])
        optimal_gbc.fit(X_train, Y_train)
        predictions = optimal_gbc.predict(X_test)
        test_acc = optimal_gbc.score(X_test, Y_test)
        test_accs.append(test_acc)
        print("test accuracy is:", test_acc)
        print(classification_report(Y_test,predictions))
        train_acc = gb.cv_results_['mean_train_score']
        val_acc = gb.cv_results_['mean_test_score']
        print("The accuarcy for cross_validation")
        plt.plot(train_acc, label = 'train_acc')
        plt.plot(val_acc, label = 'val_acc')
        plt.legend(loc = 'upper left')
        plt.show()

    print("average of test_accuracy is:", np.mean(test_accs))
```

Data1:

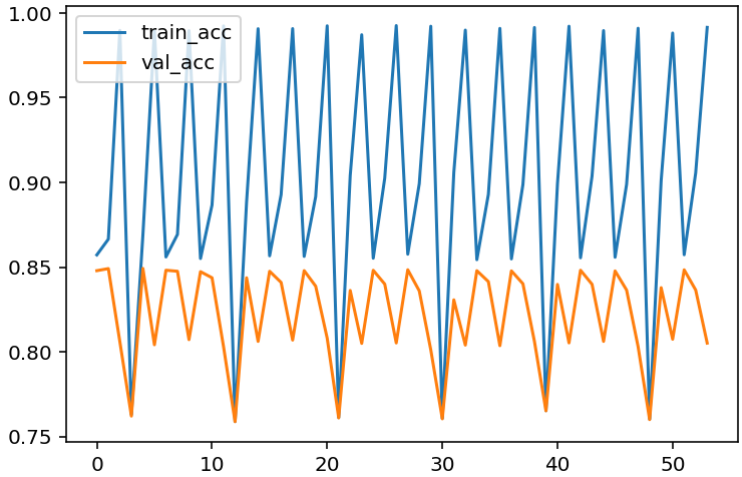
In [37]:
```
#partition1 classifier1
NN_MLP(train11_X, train11_Y, test11_X, test11_Y)
```

```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='invscaling',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='adam', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.8472286196837095
             precision    recall  f1-score   support

          0       0.87      0.93      0.90      4953
          1       0.73      0.57      0.64      1560

avg / total       0.84      0.85      0.84      6513

The accuarcy for cross_validation
```

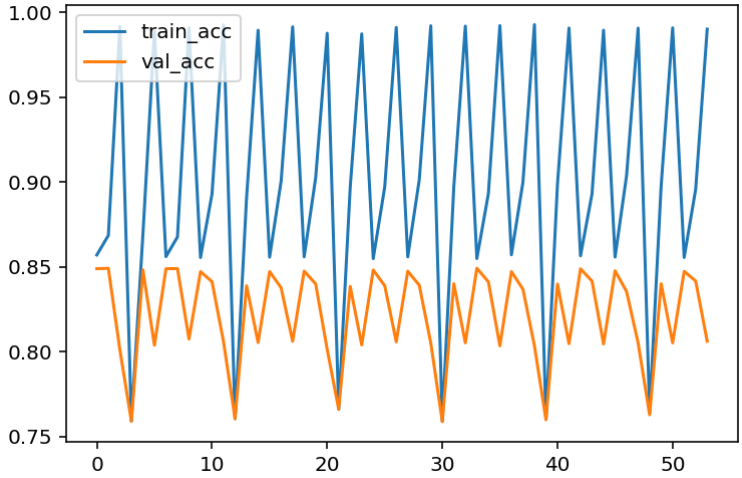



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.0001, batc
h_size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='adaptive',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.8499923230462153
             precision    recall  f1-score   support

          0       0.88      0.93      0.90      4953
          1       0.73      0.60      0.66      1560

avg / total       0.84      0.85      0.84      6513

The accuarcy for cross_validation
```

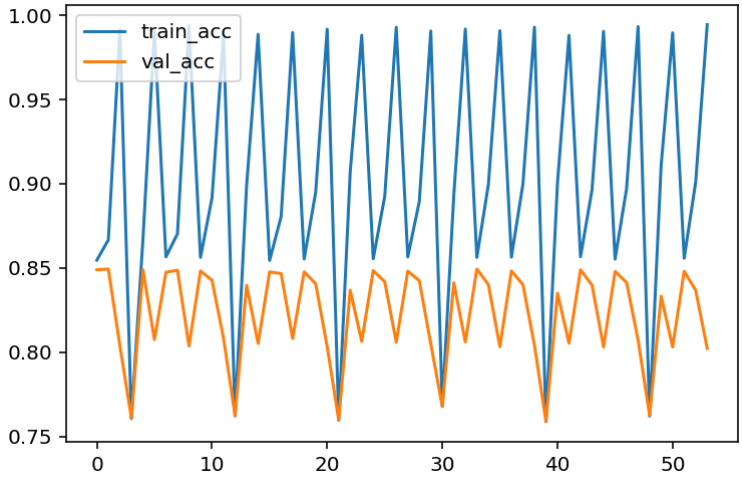



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='adam', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.8504529402732995
             precision    recall  f1-score   support

          0       0.88      0.93      0.90      4953
          1       0.73      0.59      0.65      1560

avg / total       0.84      0.85      0.84      6513

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8492246276677414
```

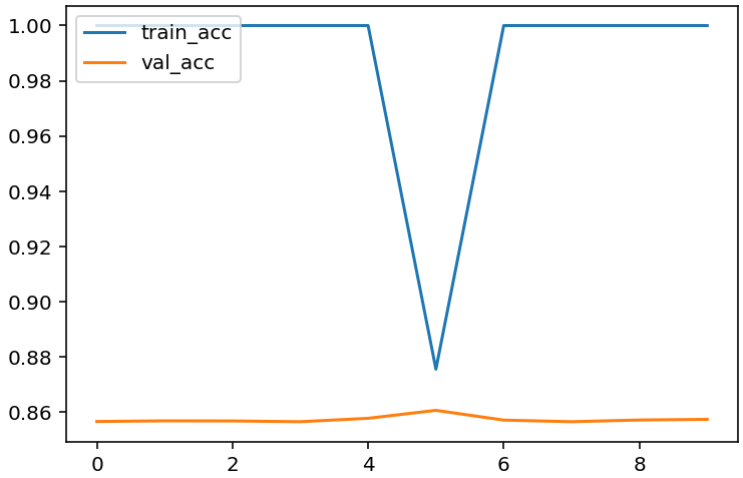
In [36]:
```
#partition1 classifier2
random_forest_classifier(train11_X, train11_Y, test11_X, test11_Y)
```

```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=10, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=200, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8556732688469215
             precision    recall  f1-score   support

          0       0.87      0.96      0.91      4953
          1       0.79      0.54      0.64      1560

avg / total       0.85      0.86      0.85      6513

The accuarcy for cross_validation
```

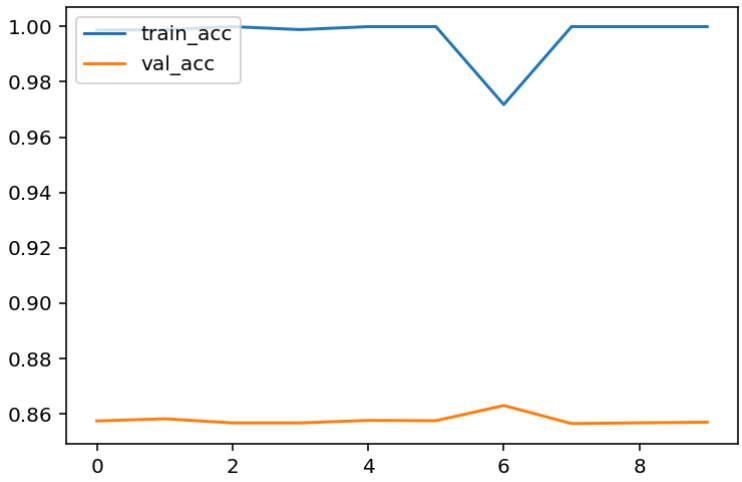



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=20, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1800, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8604329801934593
             precision    recall  f1-score   support

          0       0.88      0.94      0.91      4953
          1       0.76      0.61      0.68      1560

avg / total       0.85      0.86      0.85      6513

The accuarcy for cross_validation
```

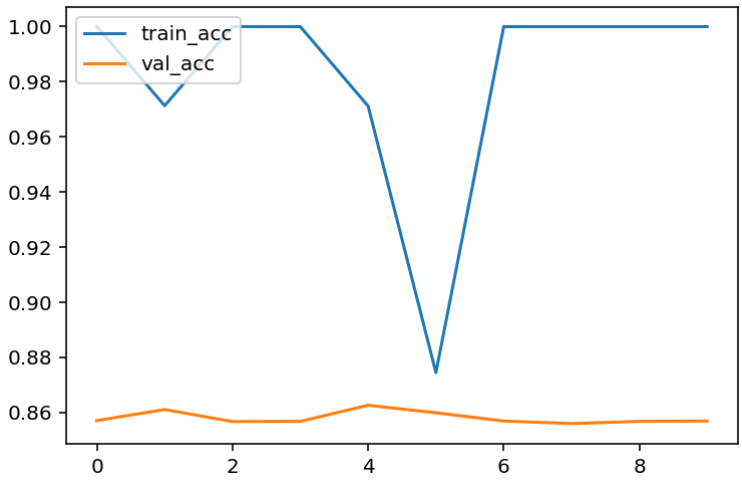



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=20, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1600, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8596652848149854
             precision    recall  f1-score   support

          0       0.88      0.94      0.91      4953
          1       0.76      0.61      0.67      1560

avg / total       0.85      0.86      0.85      6513

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8585905112851221
```

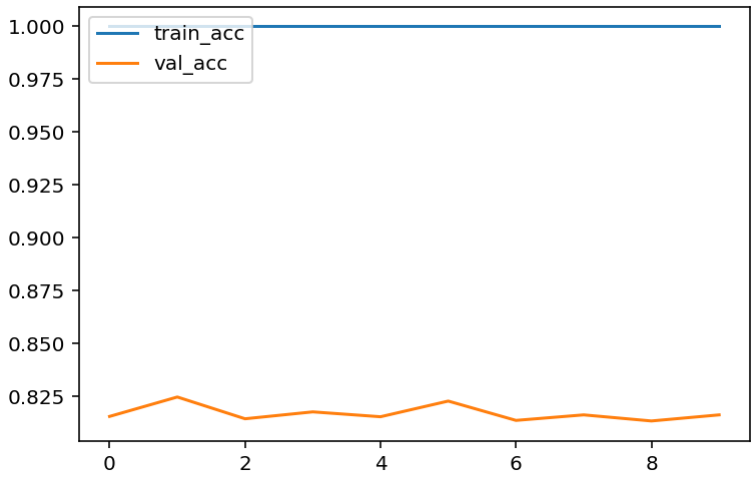
In [35]:
```
#partition1 classifier3
gb_classifier(train11_X, train11_Y, test11_X, test11_Y)
```

```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.15, loss='deviance', max_depth=110,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=2000,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8268079226163059
             precision    recall  f1-score   support

          0       0.88      0.90      0.89      4953
          1       0.65      0.60      0.62      1560

avg / total       0.82      0.83      0.82      6513

The accuarcy for cross_validation
```

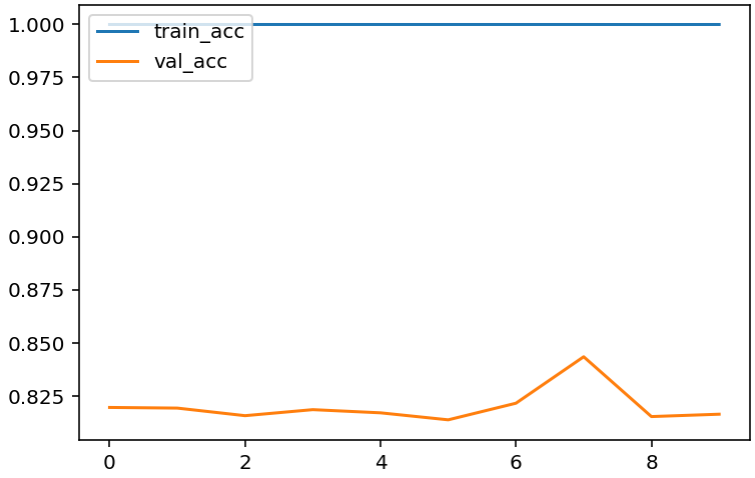



```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.15, loss='deviance', max_depth=20,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1000,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8504529402732995
             precision    recall  f1-score   support

          0       0.88      0.93      0.90      4953
          1       0.72      0.61      0.66      1560

avg / total       0.84      0.85      0.85      6513

The accuarcy for cross_validation
```

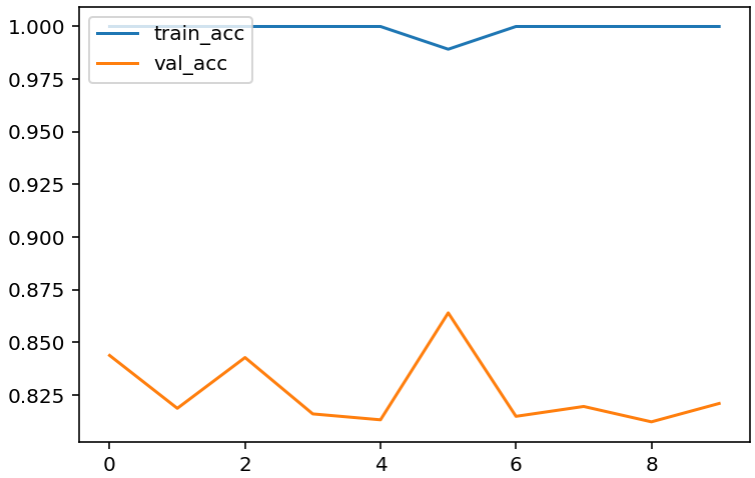



```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.01, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1600,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8687240902809765
             precision    recall  f1-score   support

          0       0.89      0.94      0.92      4953
          1       0.77      0.65      0.70      1560

avg / total       0.86      0.87      0.86      6513

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8486616510568608
```

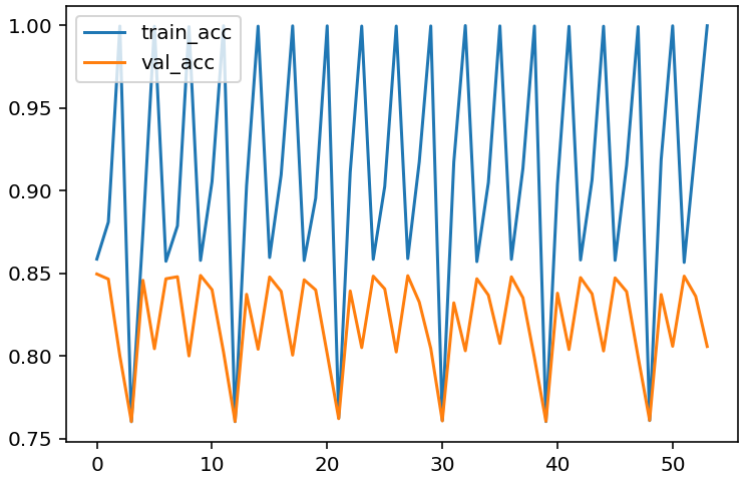
In [34]:
```
#partition2 classifier1
NN_MLP(train12_X, train12_Y, test12_X, test12_Y)
```

```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.845095510103802
             precision    recall  f1-score   support

          0       0.87      0.93      0.90     12344
          1       0.73      0.58      0.64      3937

avg / total       0.84      0.85      0.84     16281

The accuarcy for cross_validation
```

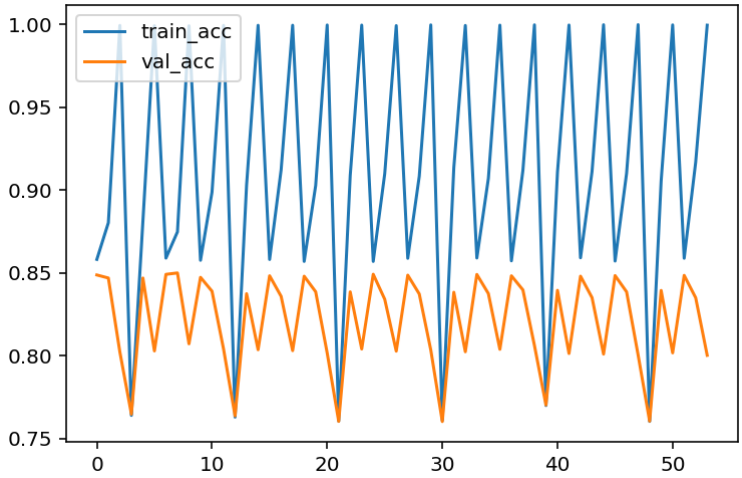



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='adaptive',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='adam', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.8449726675265647
             precision    recall  f1-score   support

          0       0.87      0.93      0.90     12344
          1       0.72      0.58      0.64      3937

avg / total       0.84      0.84      0.84     16281

The accuarcy for cross_validation
```

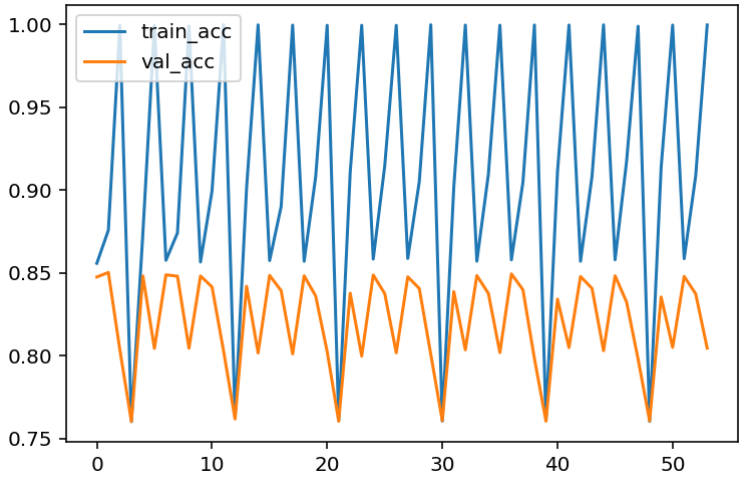



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='adam', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.8459554081444629
             precision    recall  f1-score   support

          0       0.88      0.93      0.90     12344
          1       0.72      0.59      0.65      3937

avg / total       0.84      0.85      0.84     16281

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8453411952582766
```

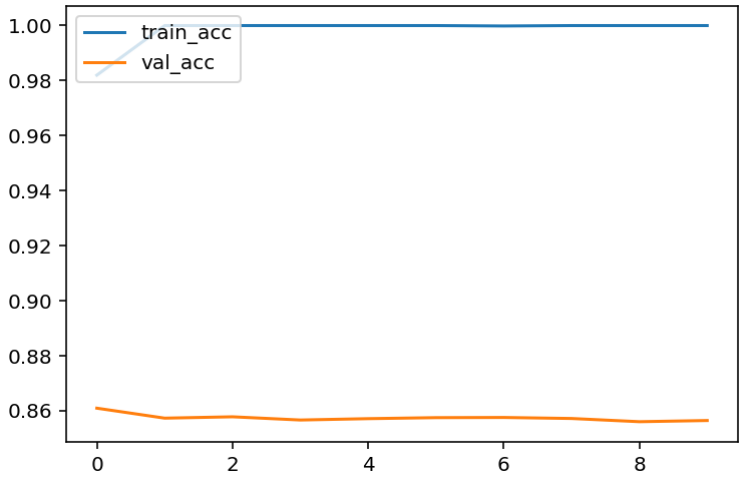
In [33]:
```
#partition2 classifier2
random_forest_classifier(train12_X, train12_Y, test12_X, test12_Y)
```

```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=20, max_features='sqrt', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=200, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8579939807137154
             precision    recall  f1-score   support

          0       0.88      0.94      0.91     12344
          1       0.76      0.60      0.67      3937

avg / total       0.85      0.86      0.85     16281

The accuarcy for cross_validation
```

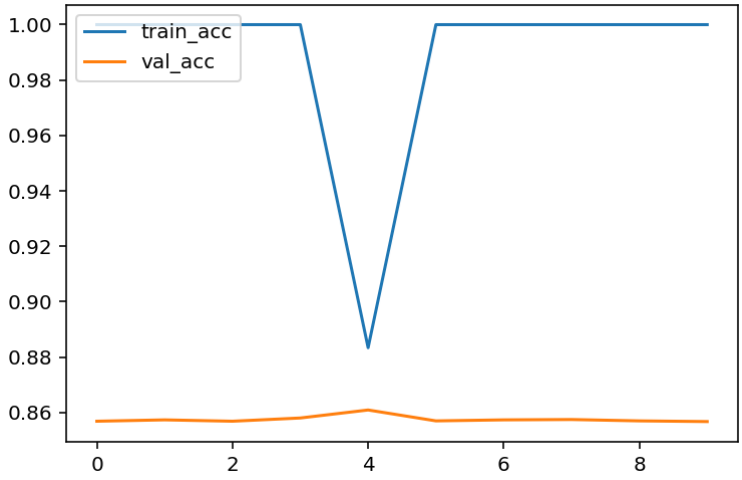



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=10, max_features='sqrt', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1800, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8583010871568085
             precision    recall  f1-score   support

          0       0.87      0.95      0.91     12344
          1       0.80      0.56      0.66      3937

avg / total       0.85      0.86      0.85     16281

The accuarcy for cross_validation
```

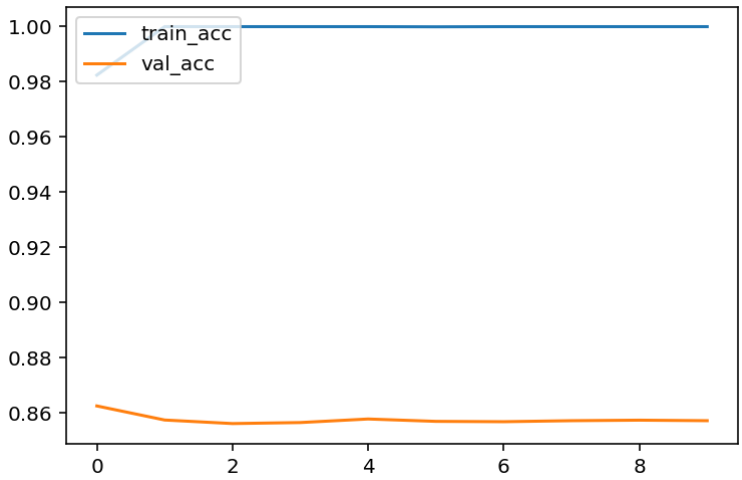



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=20, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1400, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8581782445795713
             precision    recall  f1-score   support

          0       0.88      0.94      0.91     12344
          1       0.76      0.60      0.67      3937

avg / total       0.85      0.86      0.85     16281

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8581577708166984
```

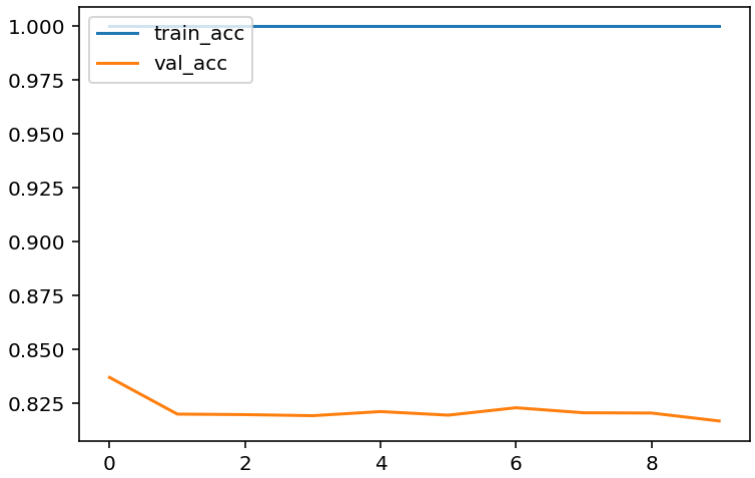
In [32]:
```
#partition2 classifier3
gb_classifier(train12_X, train12_Y, test12_X, test12_Y)
```

```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.01, loss='deviance', max_depth=20,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1400,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8356366316565321
             precision    recall  f1-score   support

          0       0.88      0.91      0.89     12344
          1       0.68      0.62      0.64      3937

avg / total       0.83      0.84      0.83     16281

The accuarcy for cross_validation
```

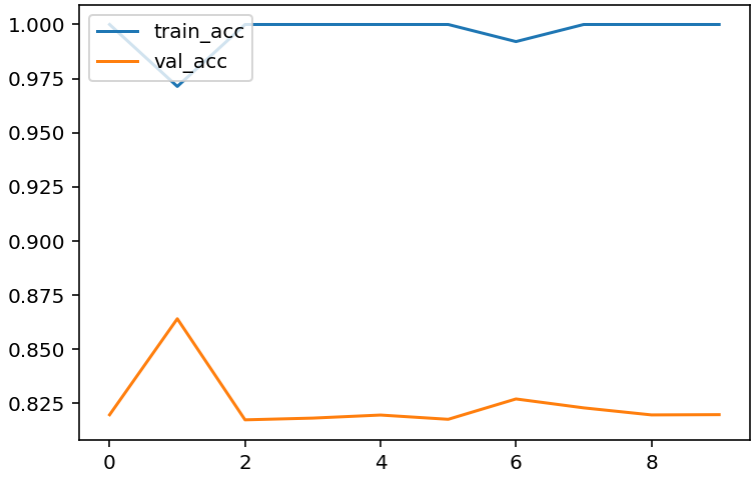



```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.01, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=800,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8618635218966894
             precision    recall  f1-score   support

          0       0.89      0.94      0.91     12344
          1       0.76      0.63      0.69      3937

avg / total       0.86      0.86      0.86     16281

The accuarcy for cross_validation
```

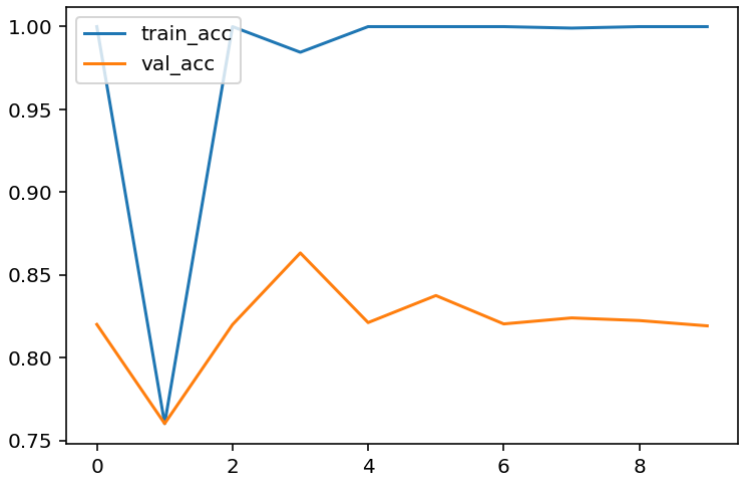



```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.01, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1000,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8635833179780111
             precision    recall  f1-score   support

          0       0.89      0.94      0.91     12344
          1       0.76      0.63      0.69      3937

avg / total       0.86      0.86      0.86     16281

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8536944905104109
```

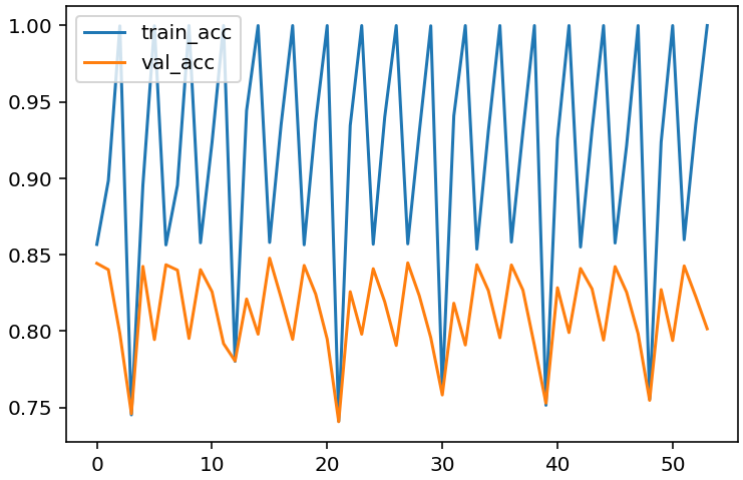
In [31]:
```
#partition3 classifier1
NN_MLP(train13_X, train13_Y, test13_X, test13_Y)
```

```
best estimator is:  MLPClassifier(activation='relu', alpha=0.01, batch_
size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='adaptive',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.8466351875311913
             precision    recall  f1-score   support

          0       0.88      0.93      0.90     19802
          1       0.72      0.59      0.65      6247

avg / total       0.84      0.85      0.84     26049

The accuarcy for cross_validation
```

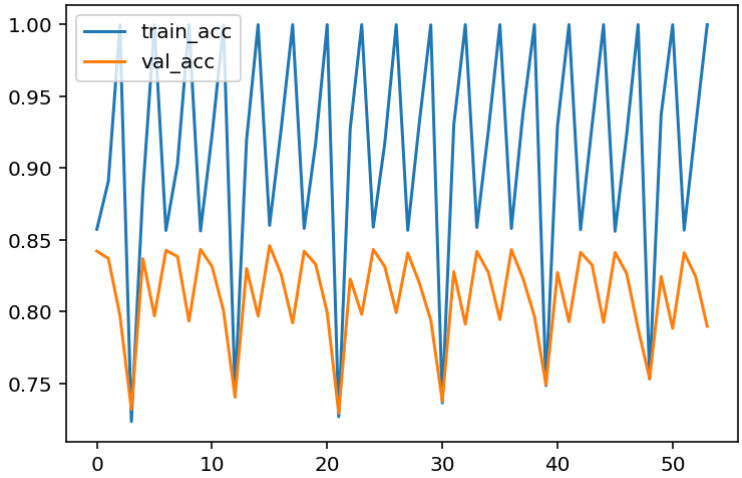



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.01, batch_
size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='adaptive',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.8455219010326692
             precision    recall  f1-score   support

          0       0.87      0.93      0.90     19802
          1       0.72      0.58      0.64      6247

avg / total       0.84      0.85      0.84     26049

The accuarcy for cross_validation
```

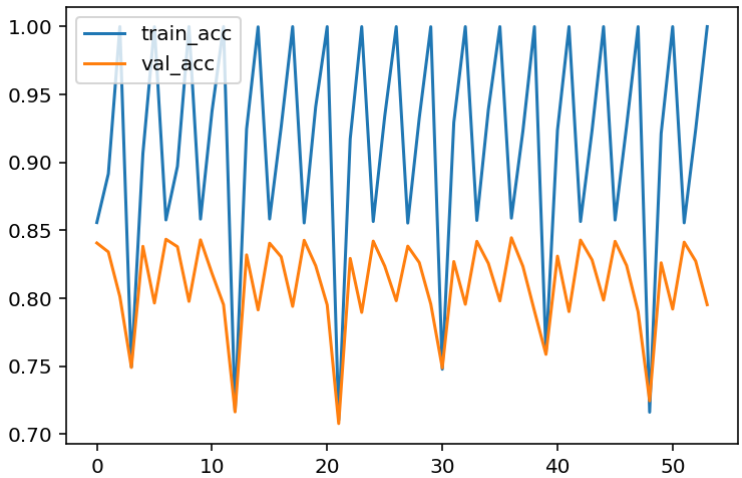



```
best estimator is:  MLPClassifier(activation='relu', alpha=1e-05, batch
_size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.8434872739836462
             precision    recall  f1-score   support

          0       0.87      0.93      0.90     19802
          1       0.72      0.58      0.64      6247

avg / total       0.84      0.84      0.84     26049

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8452147875158356
```

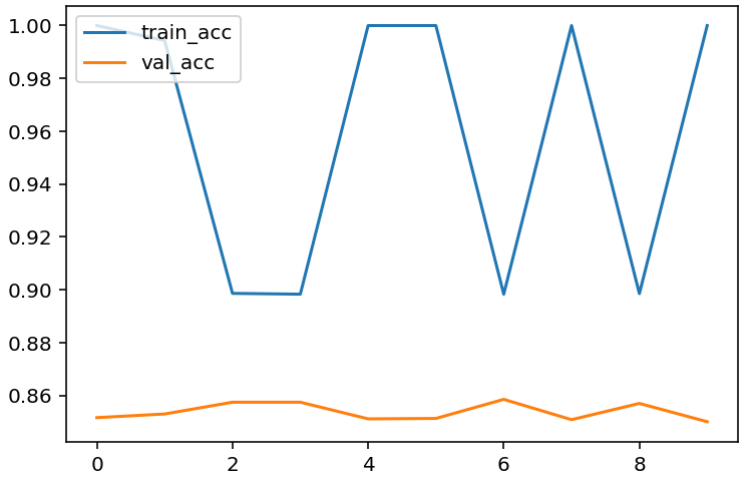
In [30]:
```
#partition3 classifier2
random_forest_classifier(train13_X, train13_Y, test13_X, test13_Y)
```

```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=10, max_features='sqrt', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=2000, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8579983876540366
             precision    recall  f1-score   support

          0       0.87      0.95      0.91     19802
          1       0.79      0.56      0.65      6247

avg / total       0.85      0.86      0.85     26049

The accuarcy for cross_validation
```

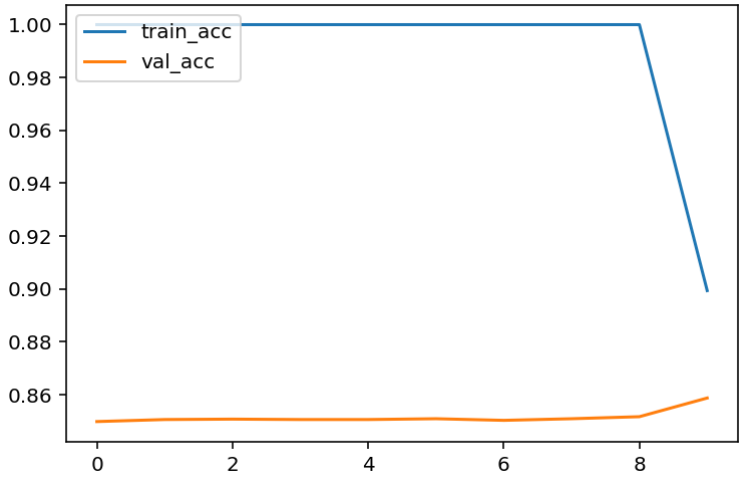



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=10, max_features='sqrt', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1400, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8580751660332451
             precision    recall  f1-score   support

          0       0.87      0.95      0.91     19802
          1       0.79      0.56      0.65      6247

avg / total       0.85      0.86      0.85     26049

The accuarcy for cross_validation
```

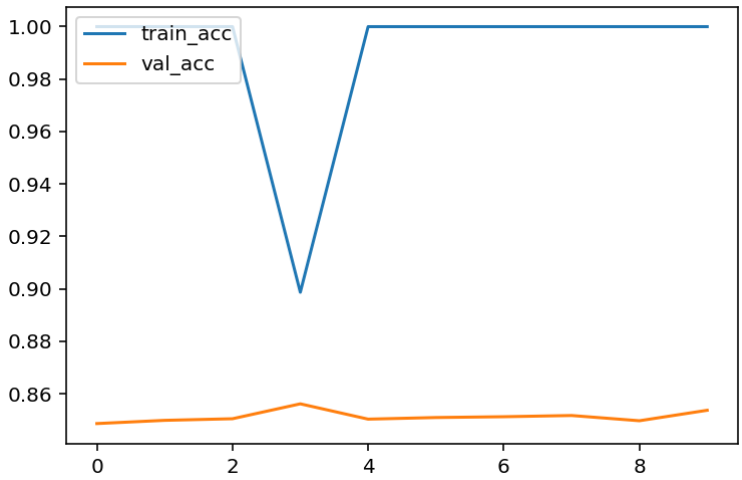



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=10, max_features='sqrt', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=600, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8576528849475987
             precision    recall  f1-score   support

          0       0.87      0.95      0.91     19802
          1       0.79      0.56      0.65      6247

avg / total       0.85      0.86      0.85     26049

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8579088128782933
```

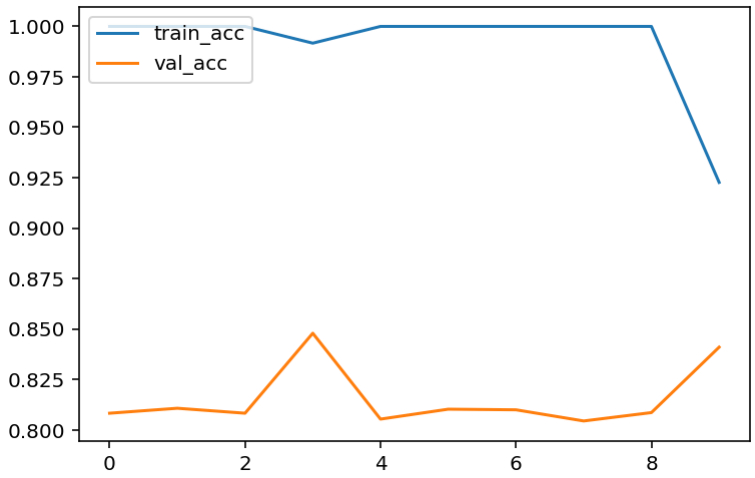
In [29]:
```
#partition2 classifier3
gb_classifier(train13_X, train13_Y, test13_X, test13_Y)
```

```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.01, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=600,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8584974471188913
             precision    recall  f1-score   support

          0       0.88      0.94      0.91     19802
          1       0.75      0.61      0.67      6247

avg / total       0.85      0.86      0.85     26049

The accuarcy for cross_validation
```

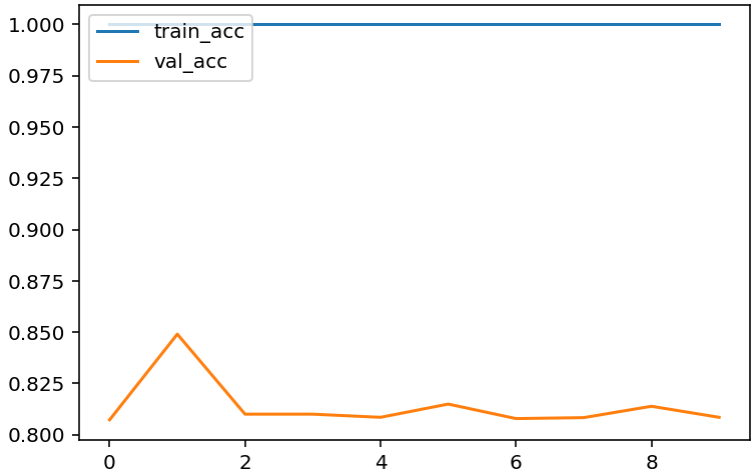



```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.15, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1600,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8544665822104496
             precision    recall  f1-score   support

          0       0.89      0.93      0.91     19802
          1       0.73      0.62      0.67      6247

avg / total       0.85      0.85      0.85     26049

The accuarcy for cross_validation
```

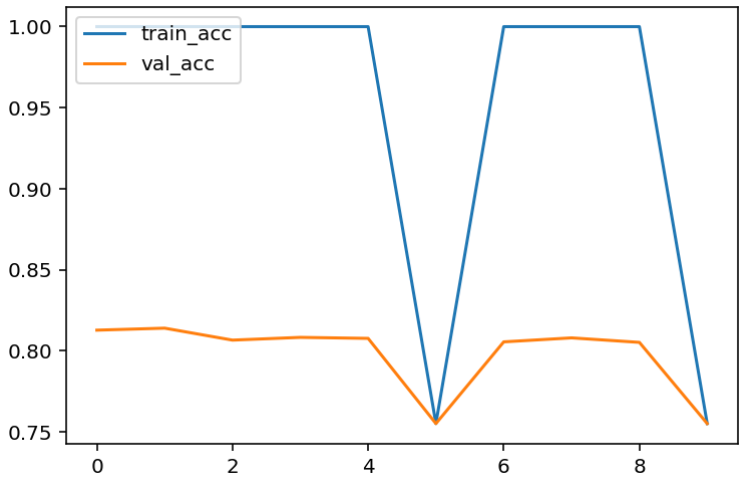



```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.15, loss='deviance', max_depth=110,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1800,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8328150792736765
             precision    recall  f1-score   support

          0       0.88      0.90      0.89     19802
          1       0.67      0.60      0.63      6247

avg / total       0.83      0.83      0.83     26049

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8485930362010058
```

Data2:

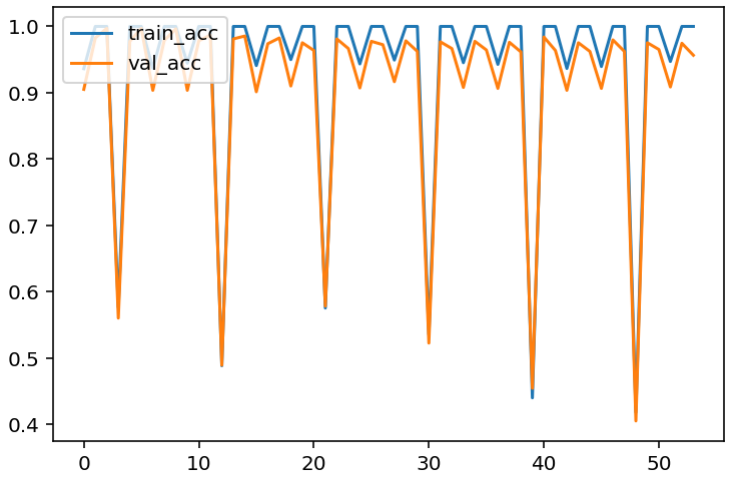
In [28]:
```
#partition1 classifier1
NN_MLP(train21_X, train21_Y, test21_X, test21_Y)
```

```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='lbfgs', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 1.0
             precision    recall  f1-score   support

          0       1.00      1.00      1.00        78
          1       1.00      1.00      1.00        12
          2       1.00      1.00      1.00       242
          3       1.00      1.00      1.00        14

avg / total       1.00      1.00      1.00       346

The accuarcy for cross_validation
```

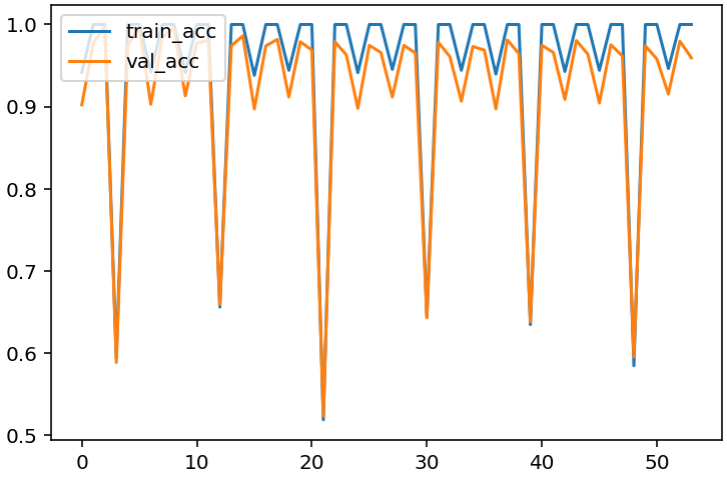



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='lbfgs', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 1.0
             precision    recall  f1-score   support

          0       1.00      1.00      1.00        78
          1       1.00      1.00      1.00        12
          2       1.00      1.00      1.00       242
          3       1.00      1.00      1.00        14

avg / total       1.00      1.00      1.00       346

The accuarcy for cross_validation
```

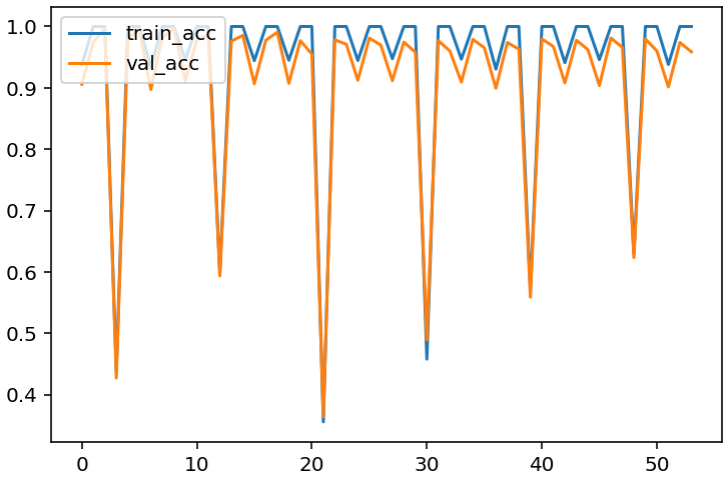



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='adaptive',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='lbfgs', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 1.0
             precision    recall  f1-score   support

          0       1.00      1.00      1.00        78
          1       1.00      1.00      1.00        12
          2       1.00      1.00      1.00       242
          3       1.00      1.00      1.00        14

avg / total       1.00      1.00      1.00       346

The accuarcy for cross_validation
```




```
average of test_accuracy is: 1.0
```

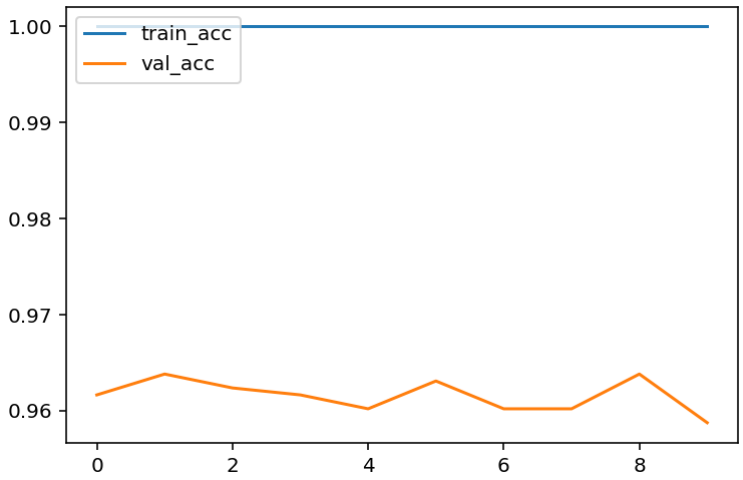
In [27]:
```
#partition1 classifier2
random_forest_classifier(train21_X, train21_Y, test21_X, test21_Y)
```

```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=90, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1600, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.976878612716763
             precision    recall  f1-score   support

          0       0.95      0.95      0.95        78
          1       0.92      0.92      0.92        12
          2       0.99      0.99      0.99       242
          3       1.00      0.93      0.96        14

avg / total       0.98      0.98      0.98       346

The accuarcy for cross_validation
```

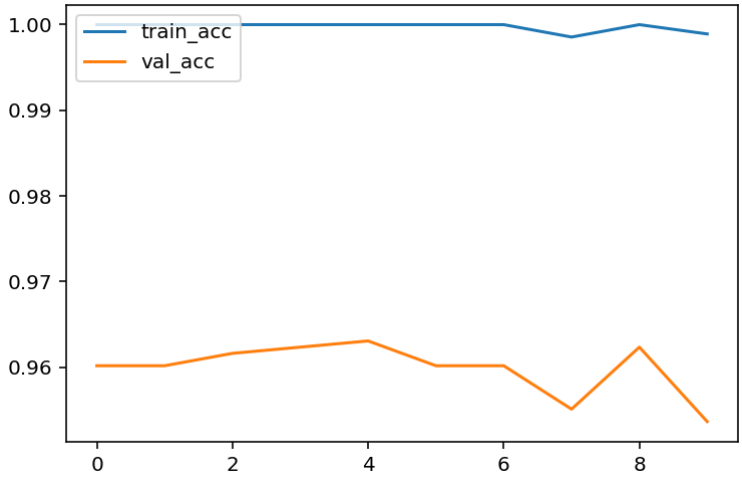



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=90, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1400, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.9797687861271677
             precision    recall  f1-score   support

          0       0.95      0.96      0.96        78
          1       1.00      0.92      0.96        12
          2       0.99      0.99      0.99       242
          3       1.00      0.93      0.96        14

avg / total       0.98      0.98      0.98       346

The accuarcy for cross_validation
```

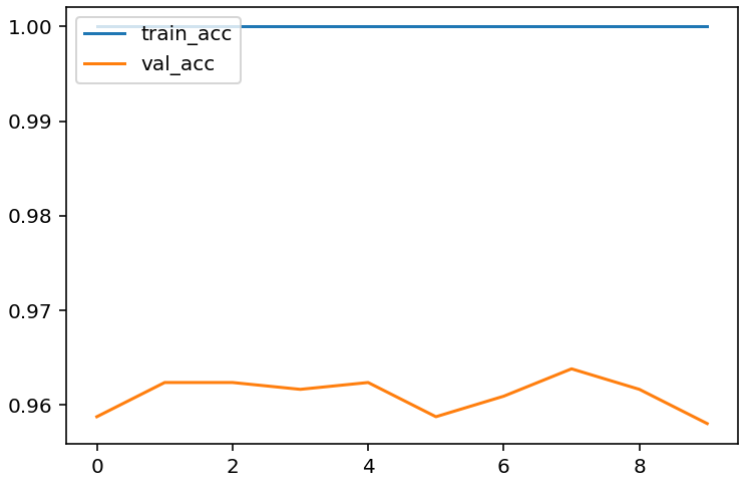



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=100, max_features='sqrt', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1000, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.9739884393063584
             precision    recall  f1-score   support

          0       0.94      0.95      0.94        78
          1       0.91      0.83      0.87        12
          2       0.99      0.99      0.99       242
          3       1.00      0.93      0.96        14

avg / total       0.97      0.97      0.97       346

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.9768786127167629
```

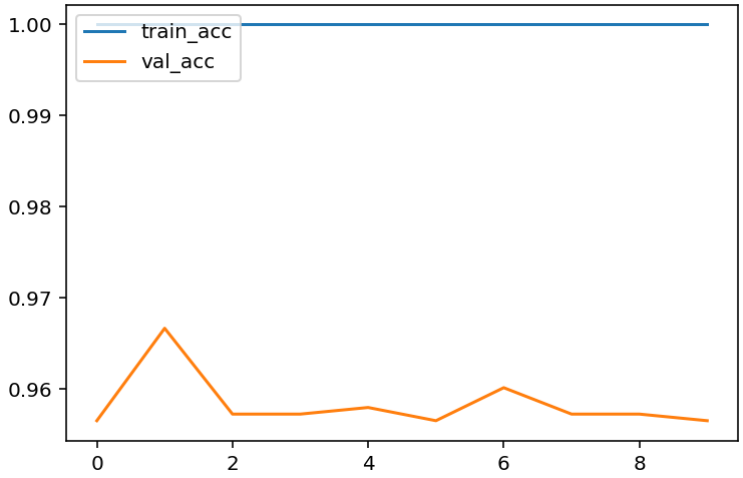
In [26]:
```
#partition1 classifier3
gb_classifier(train21_X, train21_Y, test21_X, test21_Y)
```

```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.1, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=600,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.9884393063583815
             precision    recall  f1-score   support

          0       0.99      0.96      0.97        78
          1       0.92      1.00      0.96        12
          2       0.99      1.00      0.99       242
          3       1.00      1.00      1.00        14

avg / total       0.99      0.99      0.99       346

The accuarcy for cross_validation
```

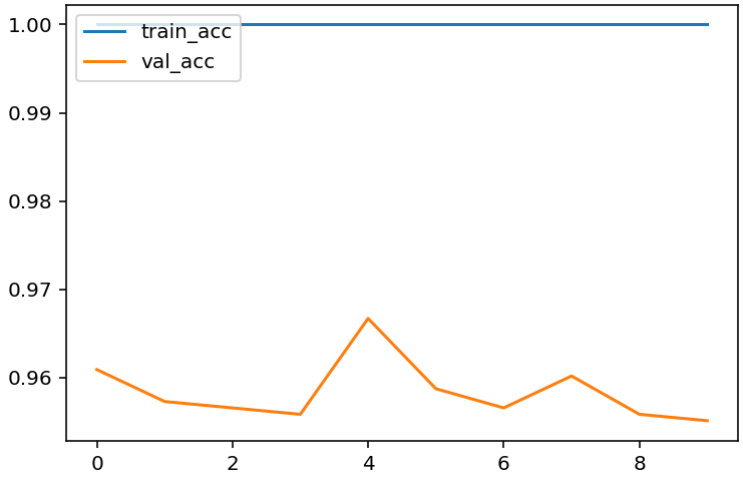


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.005, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=600,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.9826589595375722
             precision    recall  f1-score   support

          0       0.99      0.94      0.96        78
          1       0.92      1.00      0.96        12
          2       0.98      1.00      0.99       242
          3       1.00      1.00      1.00        14

avg / total       0.98      0.98      0.98       346

The accuarcy for cross_validation
```

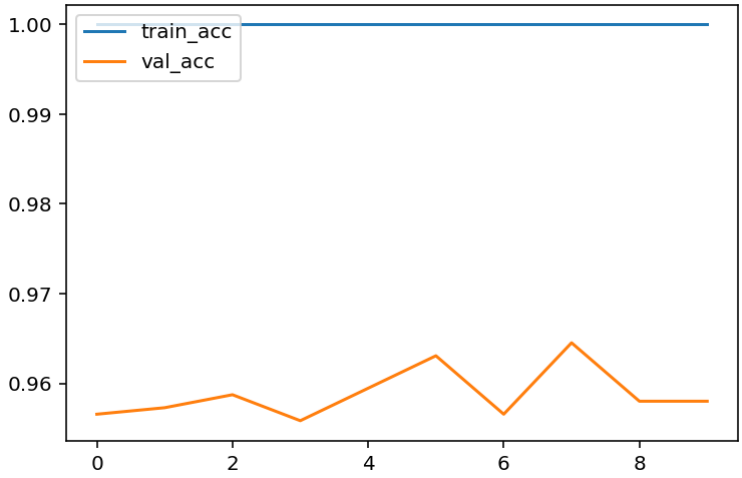


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.05, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=800,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.9884393063583815
             precision    recall  f1-score   support

          0       0.99      0.96      0.97        78
          1       0.92      1.00      0.96        12
          2       0.99      1.00      0.99       242
          3       1.00      1.00      1.00        14

avg / total       0.99      0.99      0.99       346

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.9865125240847784
```

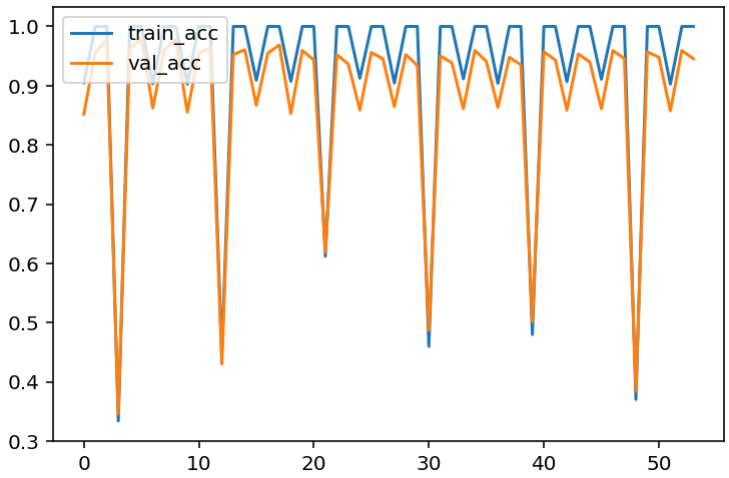
In [25]:
```
#partition2 classifier1
NN_MLP(train22_X, train22_Y, test22_X, test22_Y)
```

```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='invscaling',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='lbfgs', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.9895833333333334
             precision    recall  f1-score   support

          0       0.97      0.98      0.98       183
          1       0.97      0.92      0.94        36
          2       1.00      1.00      1.00       608
          3       1.00      0.92      0.96        37

avg / total       0.99      0.99      0.99       864

The accuarcy for cross_validation
```

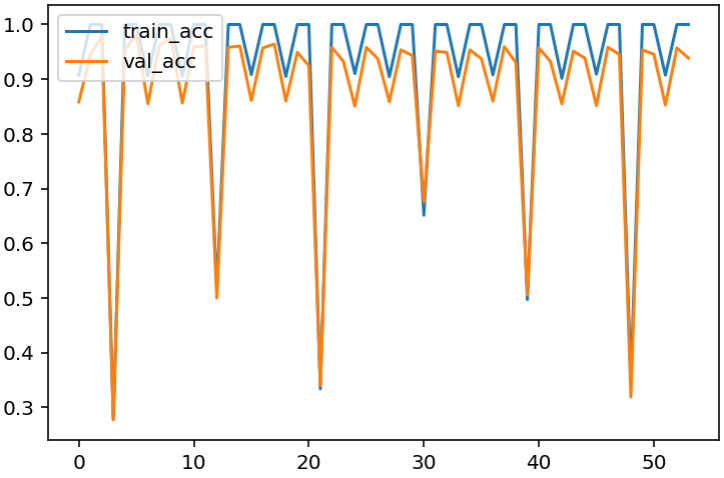



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='invscaling',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='lbfgs', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.9953703703703703
             precision    recall  f1-score   support

          0       0.98      1.00      0.99       183
          1       1.00      0.92      0.96        36
          2       1.00      1.00      1.00       608
          3       1.00      0.97      0.99        37

avg / total       1.00      1.00      1.00       864

The accuarcy for cross_validation
```

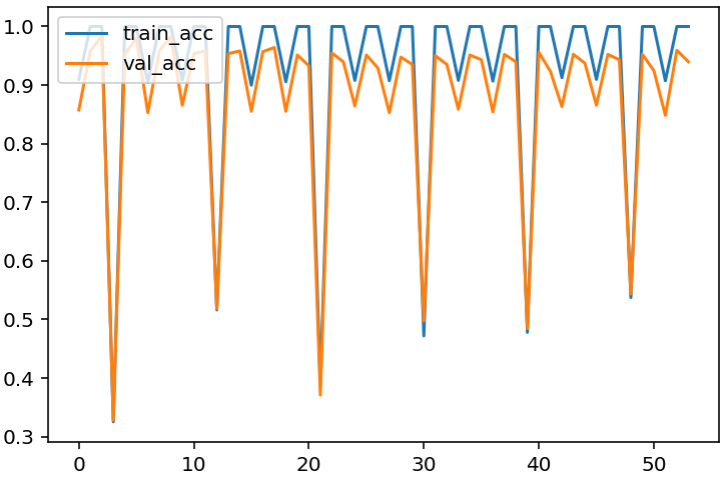



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='lbfgs', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.9895833333333334
             precision    recall  f1-score   support

          0       0.97      0.98      0.98       183
          1       1.00      0.89      0.94        36
          2       1.00      1.00      1.00       608
          3       1.00      0.95      0.97        37

avg / total       0.99      0.99      0.99       864

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.9915123456790124
```

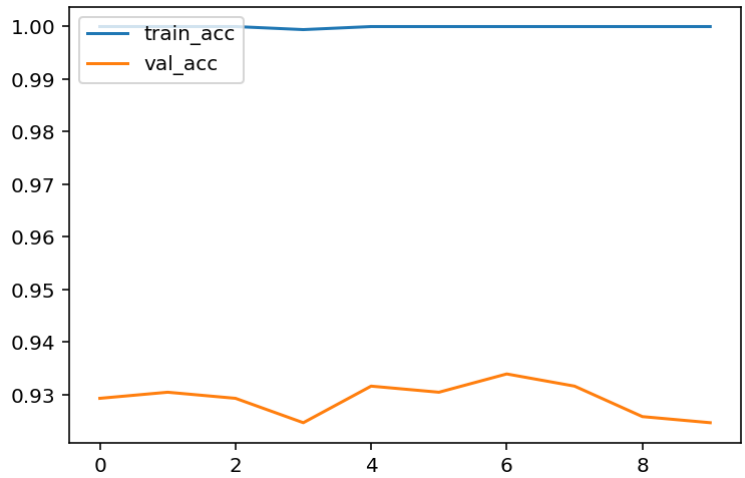
In [24]:
```
#partition2 classifier2
random_forest_classifier(train22_X, train22_Y, test22_X, test22_Y)
```

```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=10, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1200, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.9375
             precision    recall  f1-score   support

          0       0.81      0.93      0.86       183
          1       0.83      0.53      0.64        36
          2       0.98      0.97      0.98       608
          3       1.00      0.81      0.90        37

avg / total       0.94      0.94      0.94       864

The accuarcy for cross_validation
```

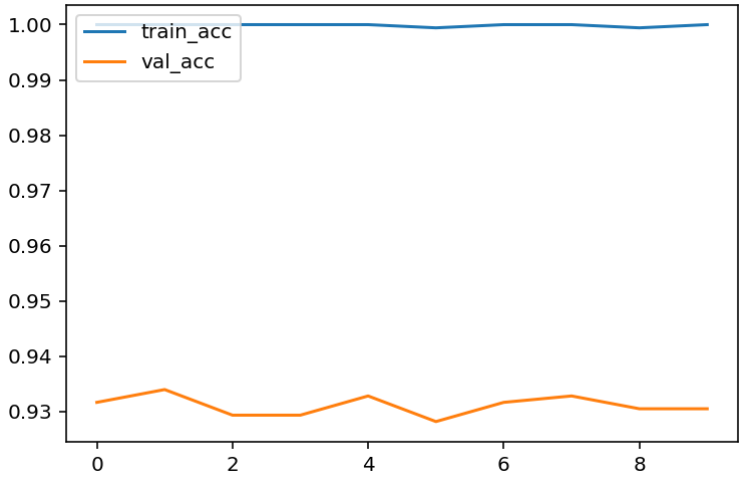



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=100, max_features='sqrt', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1200, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.9490740740740741
             precision    recall  f1-score   support

          0       0.84      0.93      0.89       183
          1       0.86      0.67      0.75        36
          2       0.99      0.97      0.98       608
          3       1.00      0.89      0.94        37

avg / total       0.95      0.95      0.95       864

The accuarcy for cross_validation
```

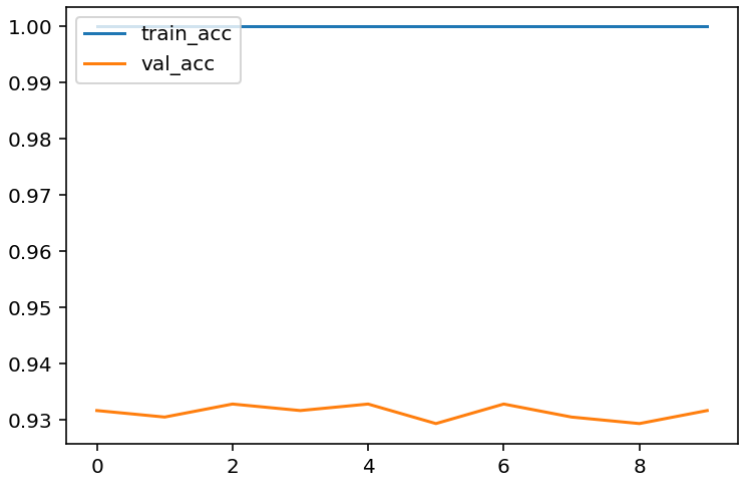



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=90, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1400, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.9444444444444444
             precision    recall  f1-score   support

          0       0.83      0.93      0.88       183
          1       0.86      0.69      0.77        36
          2       0.98      0.97      0.98       608
          3       1.00      0.81      0.90        37

avg / total       0.95      0.94      0.94       864

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.9436728395061728
```

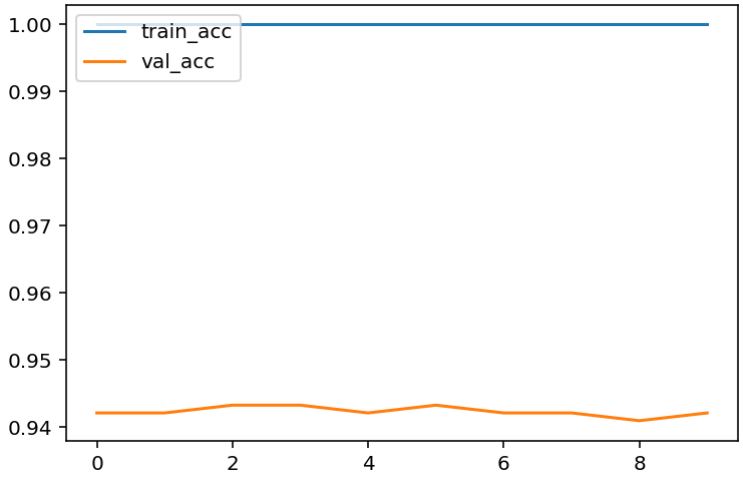
In [23]:
```
#partition2 classifier3
gb_classifier(train22_X, train22_Y, test22_X, test22_Y)
```

```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.1, loss='deviance', max_depth=70,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1600,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.9386574074074074
             precision    recall  f1-score   support

          0       0.85      0.90      0.88       183
          1       0.96      0.72      0.83        36
          2       0.96      0.98      0.97       608
          3       1.00      0.73      0.84        37

avg / total       0.94      0.94      0.94       864

The accuarcy for cross_validation
```

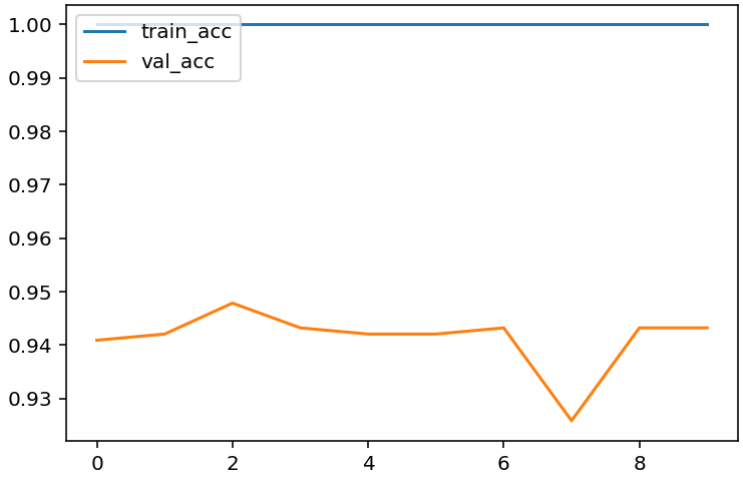


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.15, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=2000,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.9502314814814815
             precision    recall  f1-score   support

          0       0.86      0.93      0.90       183
          1       0.90      0.72      0.80        36
          2       0.98      0.98      0.98       608
          3       1.00      0.70      0.83        37

avg / total       0.95      0.95      0.95       864

The accuarcy for cross_validation
```

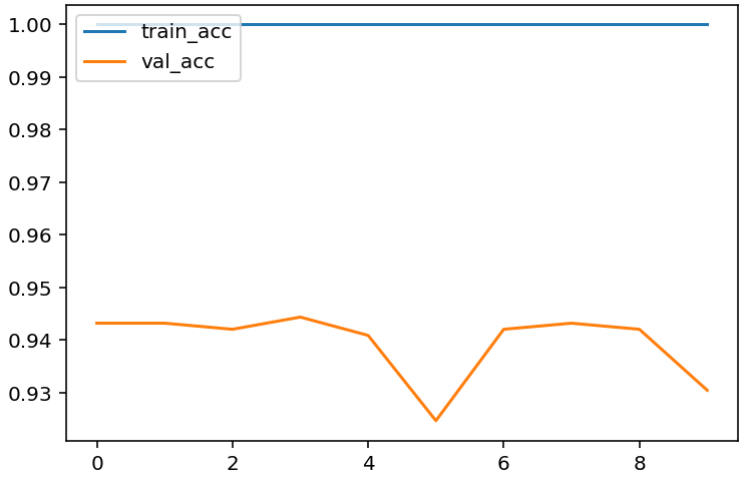


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.1, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=800,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.9502314814814815
             precision    recall  f1-score   support

          0       0.86      0.93      0.90       183
          1       0.87      0.72      0.79        36
          2       0.98      0.98      0.98       608
          3       1.00      0.70      0.83        37

avg / total       0.95      0.95      0.95       864

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.9463734567901234
```

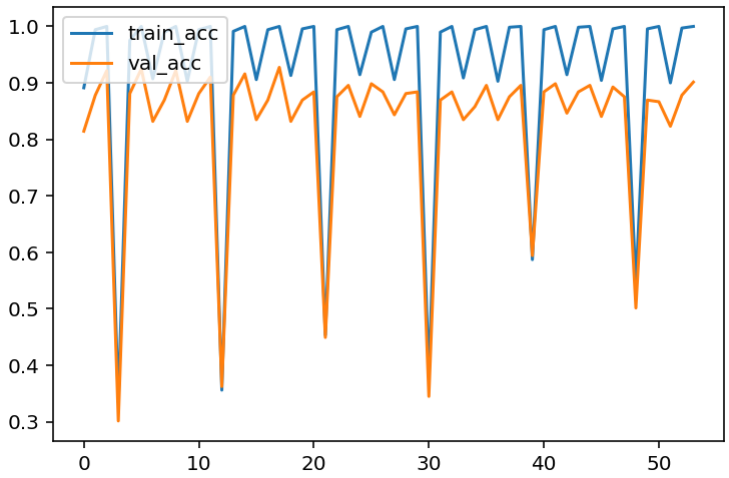
In [22]:
```
#partition3 classifier1
NN_MLP(train23_X, train23_Y, test23_X, test23_Y)
```

```
best estimator is:  MLPClassifier(activation='relu', alpha=0.01, batch_
size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='adaptive',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='lbfgs', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.9341534008683068
             precision    recall  f1-score   support

          0       0.85      0.88      0.86       300
          1       0.71      0.62      0.66        60
          2       0.98      0.97      0.98       970
          3       0.84      0.90      0.87        52

avg / total       0.93      0.93      0.93      1382

The accuarcy for cross_validation
```

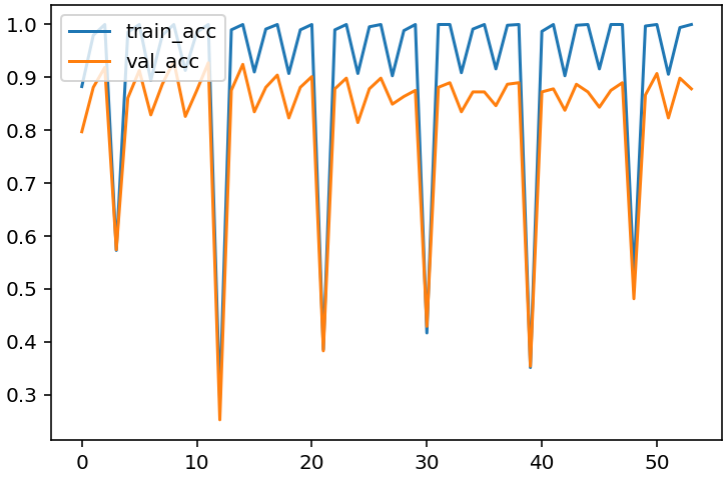



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='adaptive',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='lbfgs', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.9435600578871202
             precision    recall  f1-score   support

          0       0.86      0.89      0.88       300
          1       0.85      0.57      0.68        60
          2       0.98      0.98      0.98       970
          3       0.92      0.90      0.91        52

avg / total       0.94      0.94      0.94      1382

The accuarcy for cross_validation
```

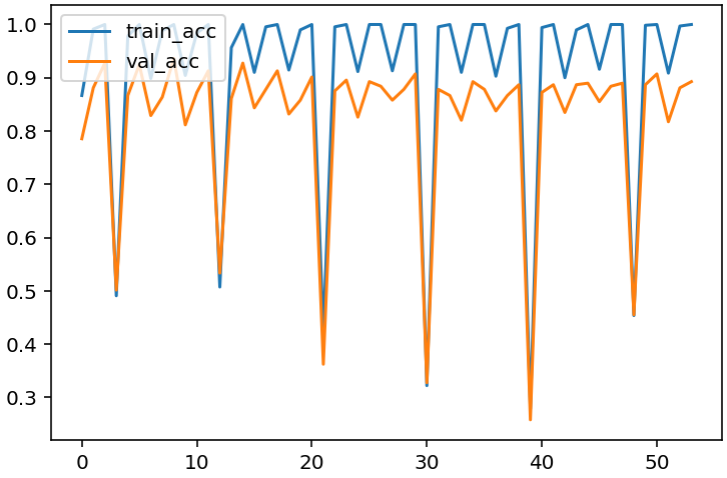



```
best estimator is:  MLPClassifier(activation='relu', alpha=0.1, batch_s
ize='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='adaptive',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='lbfgs', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.9457308248914617
             precision    recall  f1-score   support

          0       0.86      0.90      0.88       300
          1       0.89      0.53      0.67        60
          2       0.98      0.99      0.98       970
          3       0.92      0.90      0.91        52

avg / total       0.95      0.95      0.94      1382

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.9411480945489629
```

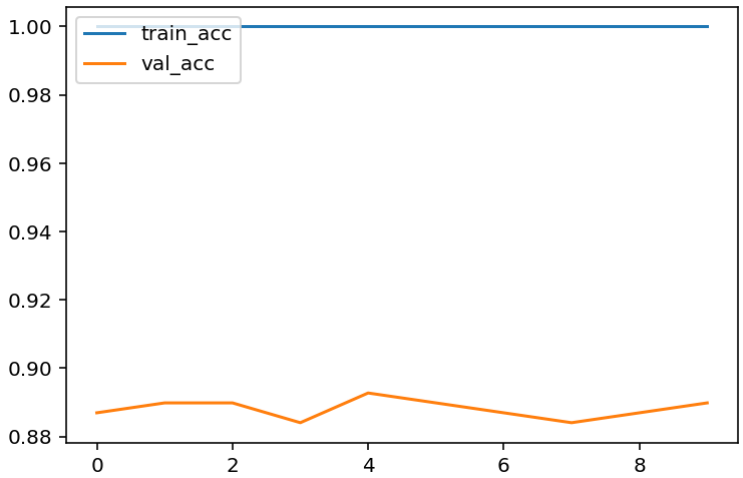
In [21]:
```
#partition3 classifier2
random_forest_classifier(train23_X, train23_Y, test23_X, test23_Y)
```

```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=110, max_features='sqrt', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=600, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8900144717800289
             precision    recall  f1-score   support

          0       0.70      0.88      0.78       300
          1       0.54      0.12      0.19        60
          2       0.97      0.95      0.96       970
          3       0.90      0.69      0.78        52

avg / total       0.89      0.89      0.88      1382

The accuarcy for cross_validation
```

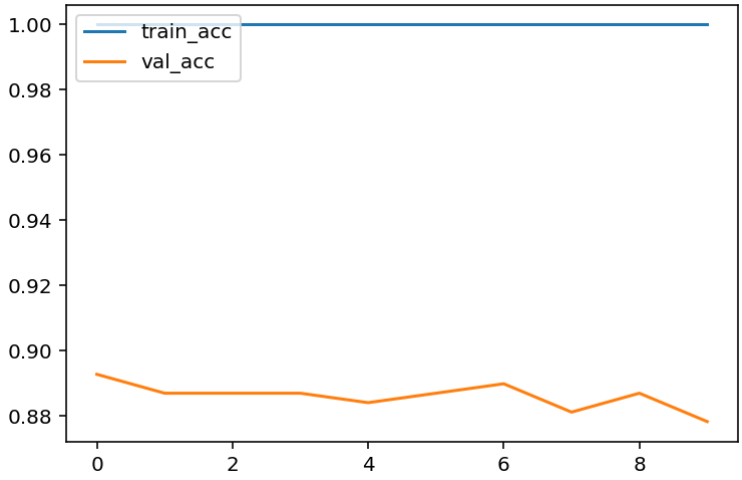



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=40, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1000, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8900144717800289
             precision    recall  f1-score   support

          0       0.70      0.88      0.78       300
          1       0.54      0.12      0.19        60
          2       0.97      0.95      0.96       970
          3       0.92      0.67      0.78        52

avg / total       0.89      0.89      0.88      1382

The accuarcy for cross_validation
```

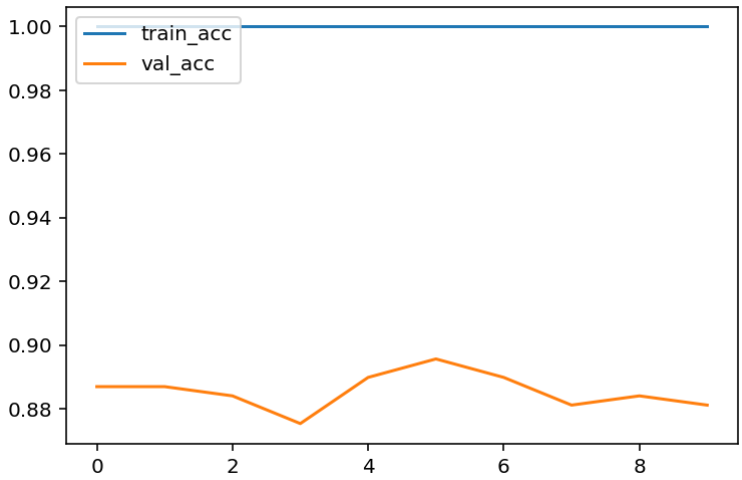



```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=90, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=800, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.8878437047756874
             precision    recall  f1-score   support

          0       0.69      0.87      0.77       300
          1       0.58      0.12      0.19        60
          2       0.97      0.95      0.96       970
          3       0.90      0.67      0.77        52

avg / total       0.89      0.89      0.88      1382

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8892908827785817
```

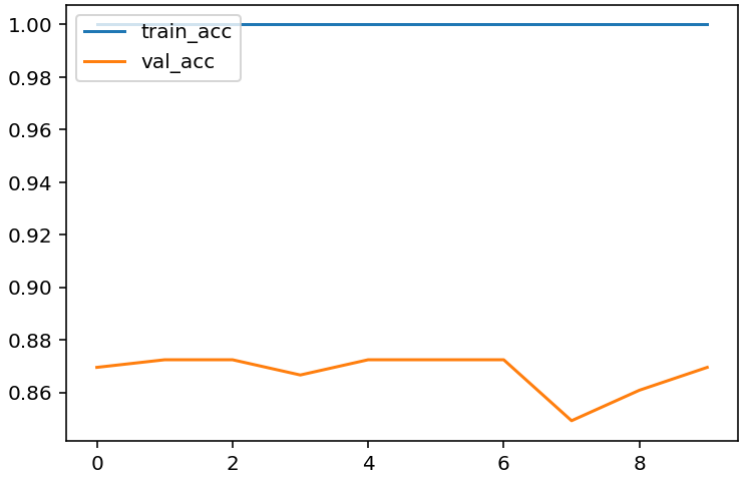
In [20]:
```
#partition3 classifier3
gb_classifier(train23_X, train23_Y, test23_X, test23_Y)
```

```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.1, loss='deviance', max_depth=30,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1000,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8668596237337193
             precision    recall  f1-score   support

          0       0.71      0.80      0.75       300
          1       0.50      0.13      0.21        60
          2       0.94      0.95      0.95       970
          3       0.58      0.50      0.54        52

avg / total       0.86      0.87      0.86      1382

The accuarcy for cross_validation
```

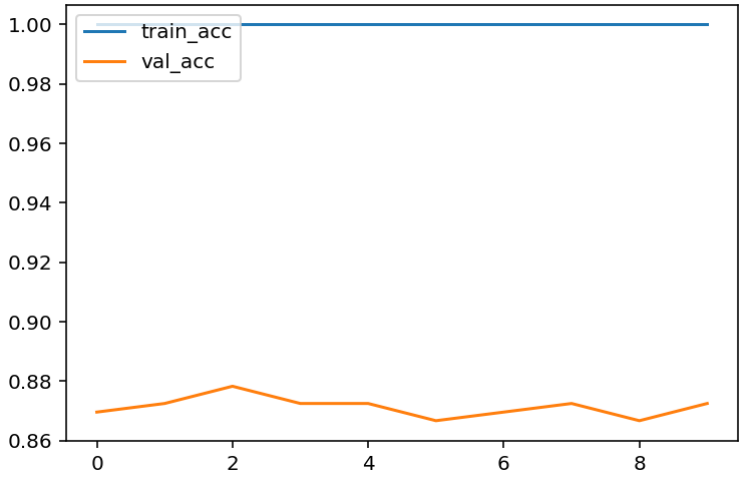


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.005, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1000,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8675832127351665
             precision    recall  f1-score   support

          0       0.71      0.80      0.75       300
          1       0.45      0.15      0.23        60
          2       0.95      0.95      0.95       970
          3       0.59      0.58      0.58        52

avg / total       0.86      0.87      0.86      1382

The accuarcy for cross_validation
```

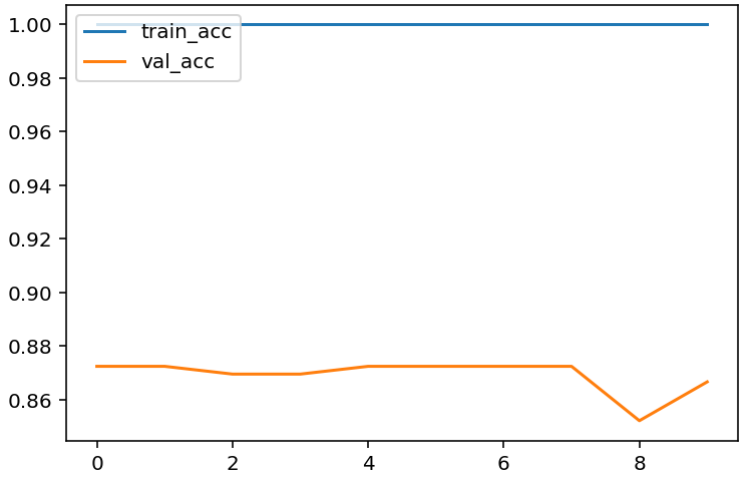


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.01, loss='deviance', max_depth=100,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1600,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.8784370477568741
             precision    recall  f1-score   support

          0       0.72      0.86      0.78       300
          1       0.46      0.10      0.16        60
          2       0.95      0.95      0.95       970
          3       0.59      0.50      0.54        52

avg / total       0.87      0.88      0.87      1382

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.8709599614085866
```

Data3

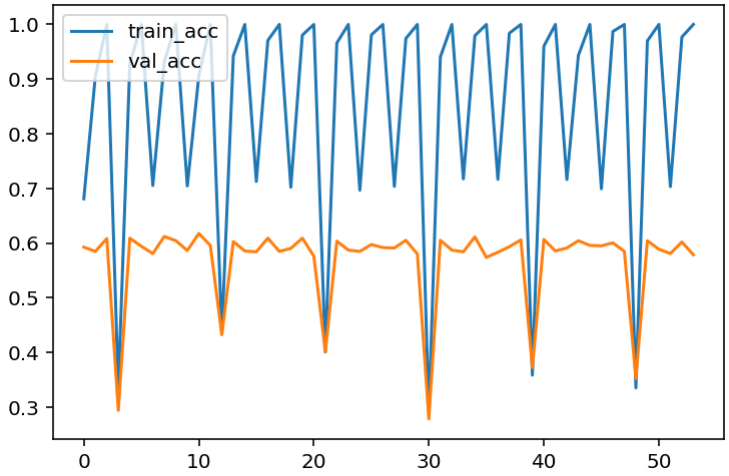
In [19]:
```
#partition1 classifier1
NN_MLP(train31_X, train31_Y, test31_X, test31_Y)
```

```
best estimator is:  MLPClassifier(activation='relu', alpha=0.01, batch_
size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='adam', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.665625
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         1
          1       0.17      0.17      0.17        12
          2       0.75      0.69      0.72       138
          3       0.67      0.74      0.70       125
          4       0.59      0.60      0.59        40
          5       0.00      0.00      0.00         4

avg / total       0.66      0.67      0.66       320

The accuarcy for cross_validation
```

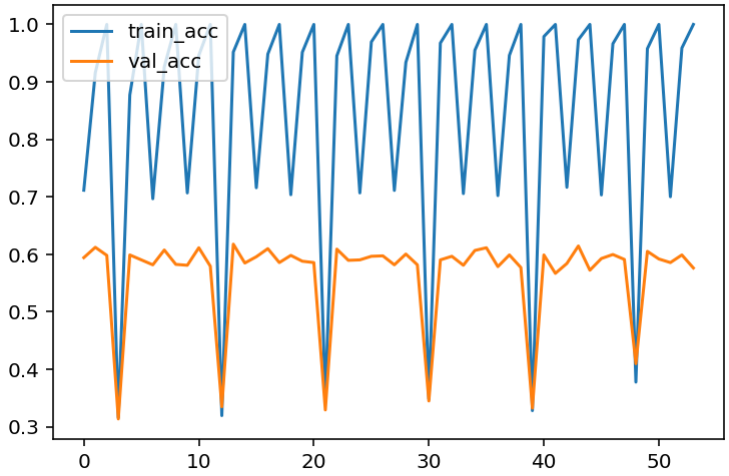


```
best estimator is:  MLPClassifier(activation='relu', alpha=0.01, batch_
size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='invscaling',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='adam', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.66875
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         1
          1       0.20      0.17      0.18        12
          2       0.72      0.79      0.75       138
          3       0.70      0.66      0.68       125
          4       0.57      0.53      0.55        40
          5       0.00      0.00      0.00         4

avg / total       0.66      0.67      0.67       320

The accuarcy for cross_validation
```

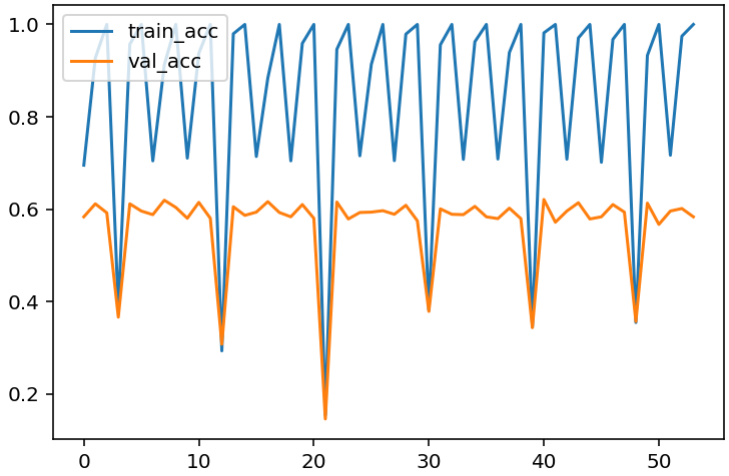


```
best estimator is:  MLPClassifier(activation='relu', alpha=1e-05, batch
_size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='invscaling',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='adam', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.675
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         1
          1       0.20      0.17      0.18        12
          2       0.75      0.74      0.74       138
          3       0.67      0.73      0.70       125
          4       0.60      0.53      0.56        40
          5       0.00      0.00      0.00         4

avg / total       0.67      0.68      0.67       320

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.6697916666666668
```

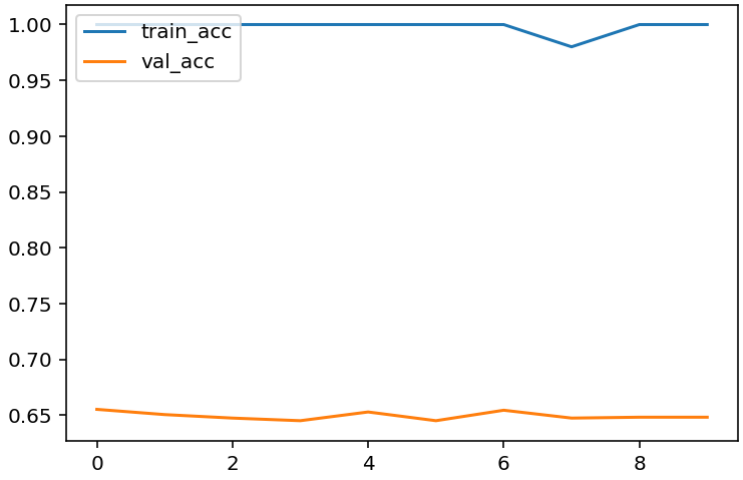
In [18]:
```
#partition1 classifier2
random_forest_classifier(train31_X, train31_Y, test31_X, test31_Y)
```

```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=60, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=400, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.728125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         1
          1       0.00      0.00      0.00        12
          2       0.74      0.86      0.80       138
          3       0.70      0.74      0.72       125
          4       0.80      0.50      0.62        40
          5       0.50      0.25      0.33         4

avg / total       0.70      0.73      0.71       320

The accuarcy for cross_validation
```

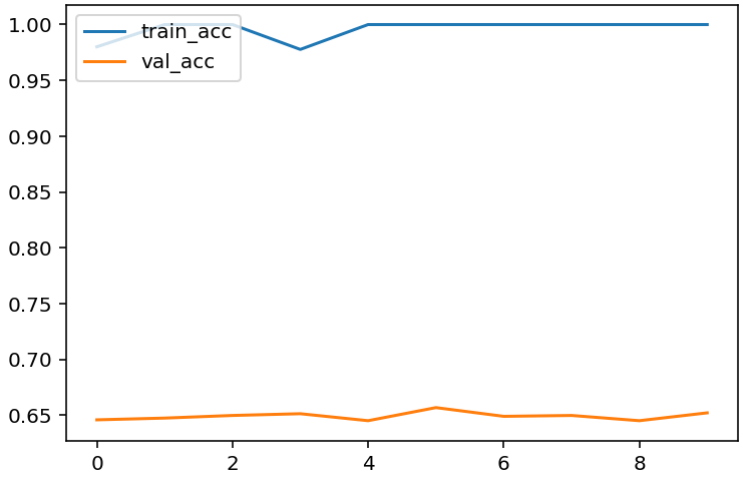


```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=40, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=600, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.721875
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         1
          1       0.00      0.00      0.00        12
          2       0.75      0.85      0.80       138
          3       0.69      0.74      0.71       125
          4       0.78      0.53      0.63        40
          5       1.00      0.25      0.40         4

avg / total       0.70      0.72      0.70       320

The accuarcy for cross_validation
```

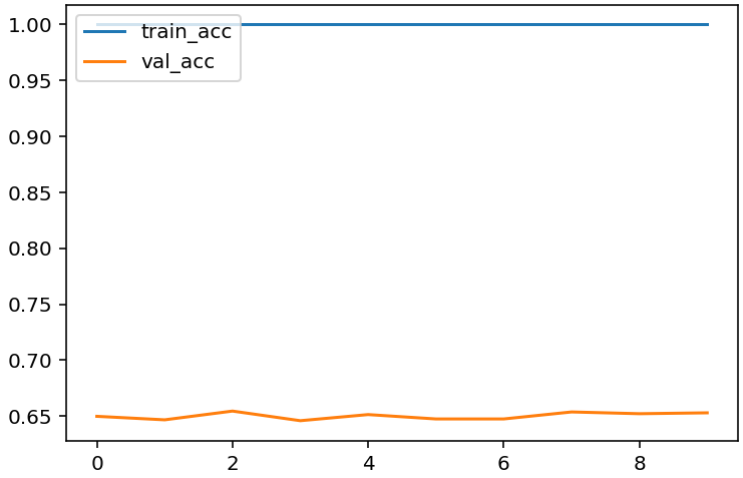


```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=90, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=800, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.73125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         1
          1       0.00      0.00      0.00        12
          2       0.76      0.86      0.81       138
          3       0.69      0.75      0.72       125
          4       0.80      0.50      0.62        40
          5       0.50      0.25      0.33         4

avg / total       0.70      0.73      0.71       320

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.7270833333333334
```

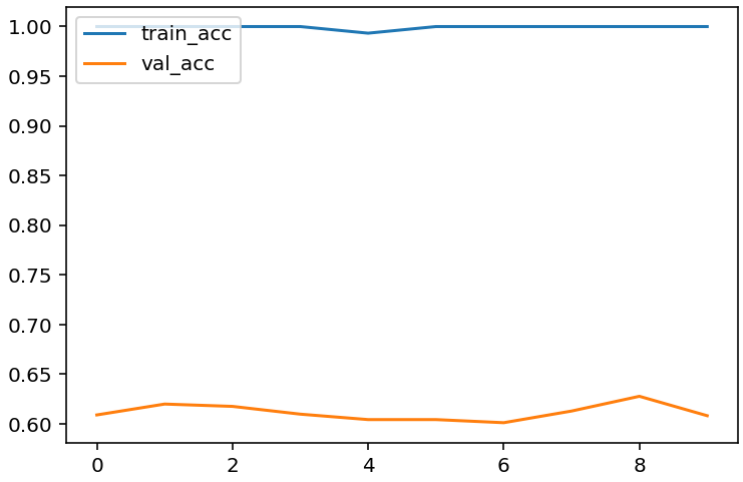
In [17]:
```
#partition1 classifier3
gb_classifier(train31_X, train31_Y, test31_X, test31_Y)
```

```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.05, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=200,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.715625
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         1
          1       0.17      0.08      0.11        12
          2       0.76      0.87      0.81       138
          3       0.70      0.70      0.70       125
          4       0.68      0.47      0.56        40
          5       0.25      0.25      0.25         4

avg / total       0.70      0.72      0.70       320

The accuarcy for cross_validation
```

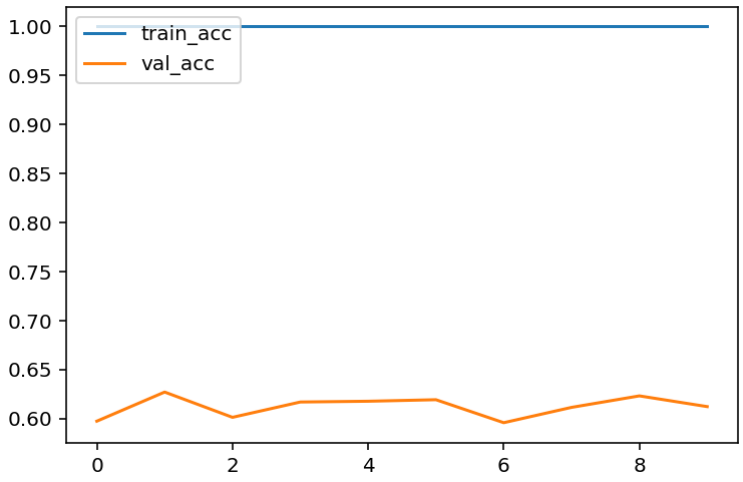


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.15, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1200,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.721875
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         1
          1       0.00      0.00      0.00        12
          2       0.76      0.84      0.80       138
          3       0.69      0.74      0.72       125
          4       0.75      0.53      0.62        40
          5       0.50      0.25      0.33         4

avg / total       0.70      0.72      0.71       320

The accuarcy for cross_validation
```

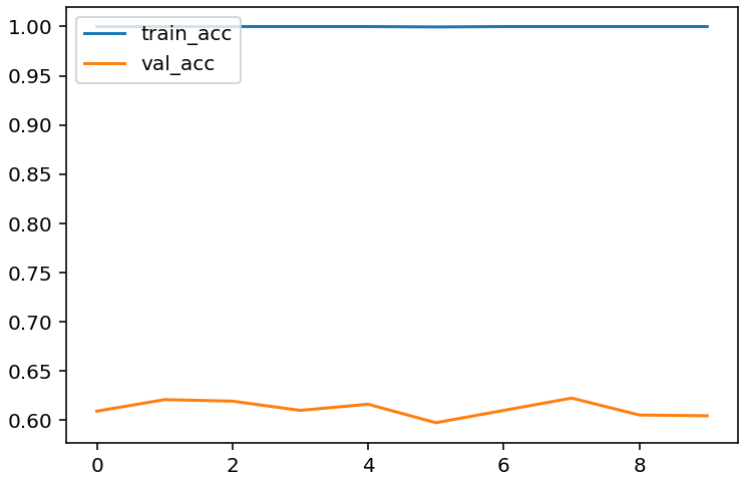


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.01, loss='deviance', max_depth=50,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=2000,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.653125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         1
          1       0.50      0.08      0.14        12
          2       0.69      0.78      0.73       138
          3       0.64      0.62      0.63       125
          4       0.56      0.55      0.56        40
          5       0.33      0.25      0.29         4

avg / total       0.64      0.65      0.64       320

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.696875
```

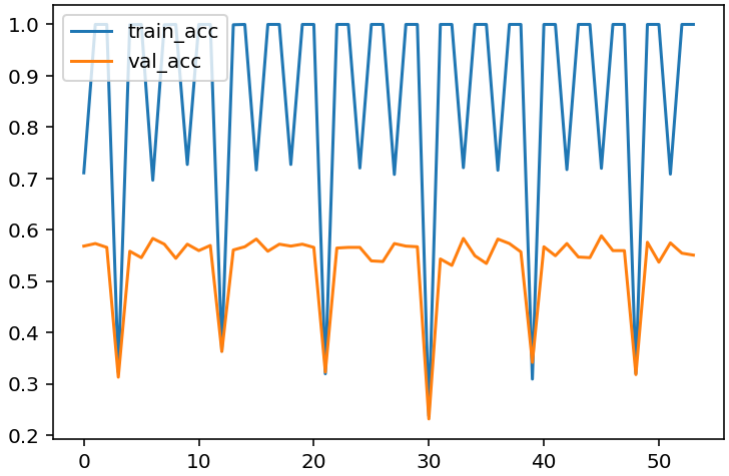
In [16]:
```
#partition2 classifier1
NN_MLP(train32_X, train32_Y, test32_X, test32_Y)
```

```
best estimator is:  MLPClassifier(activation='relu', alpha=1e-06, batch
_size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.5875
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         5
          1       0.33      0.03      0.06        29
          2       0.64      0.74      0.69       334
          3       0.56      0.56      0.56       327
          4       0.45      0.40      0.42        94
          5       0.00      0.00      0.00        11

avg / total       0.56      0.59      0.57       800

The accuarcy for cross_validation
```

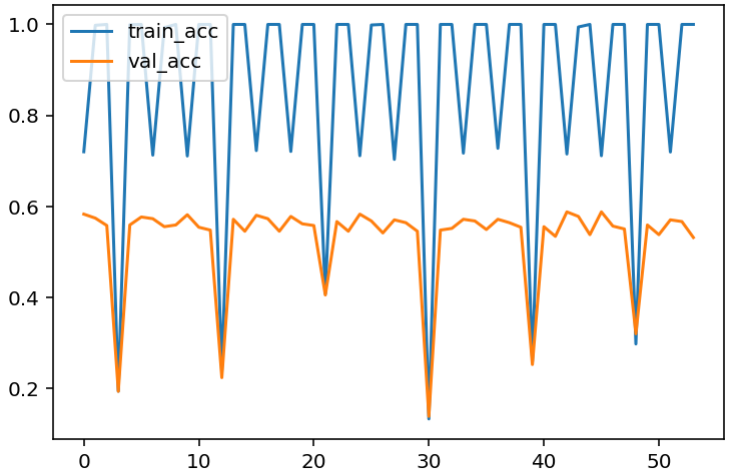


```
best estimator is:  MLPClassifier(activation='relu', alpha=1e-05, batch
_size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='adaptive',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.585
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         5
          1       0.00      0.00      0.00        29
          2       0.64      0.74      0.69       334
          3       0.57      0.55      0.56       327
          4       0.43      0.43      0.43        94
          5       0.00      0.00      0.00        11

avg / total       0.55      0.58      0.57       800

The accuarcy for cross_validation
```

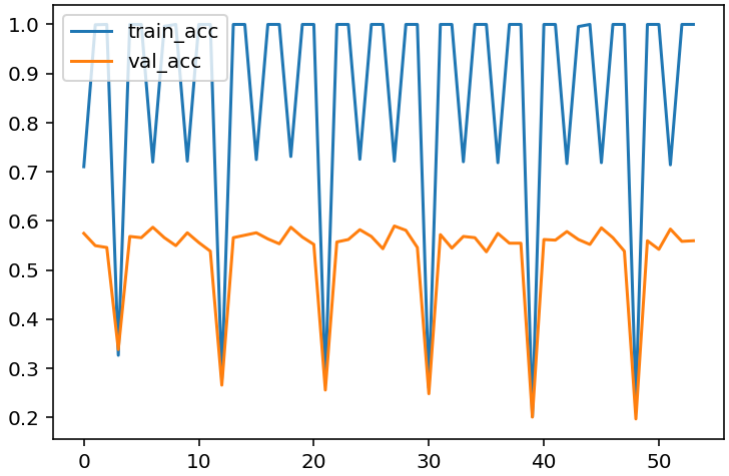


```
best estimator is:  MLPClassifier(activation='relu', alpha=0.0001, batc
h_size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.57875
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         5
          1       0.00      0.00      0.00        29
          2       0.64      0.78      0.70       334
          3       0.57      0.52      0.54       327
          4       0.37      0.35      0.36        94
          5       0.00      0.00      0.00        11

avg / total       0.54      0.58      0.56       800

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.5837499999999999
```

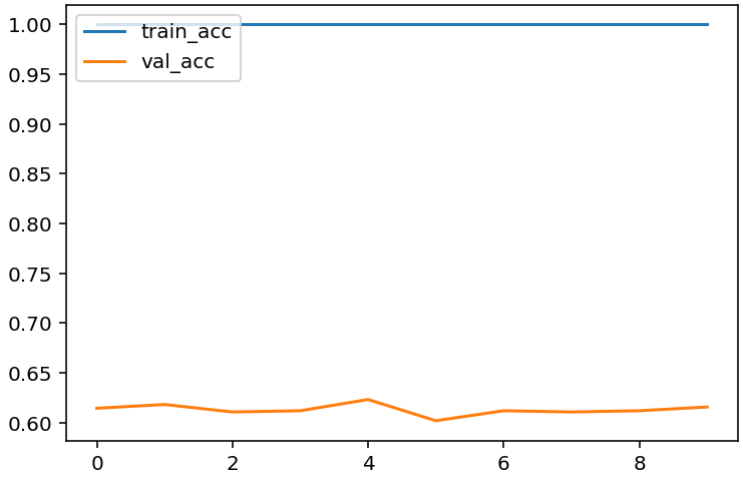
In [15]:
```
#partition2 classifier2
random_forest_classifier(train32_X, train32_Y, test32_X, test32_Y)
```

```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=90, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=200, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.665
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         5
          1       0.00      0.00      0.00        29
          2       0.71      0.81      0.75       334
          3       0.66      0.64      0.65       327
          4       0.54      0.56      0.55        94
          5       0.00      0.00      0.00        11

avg / total       0.63      0.67      0.65       800

The accuarcy for cross_validation
```

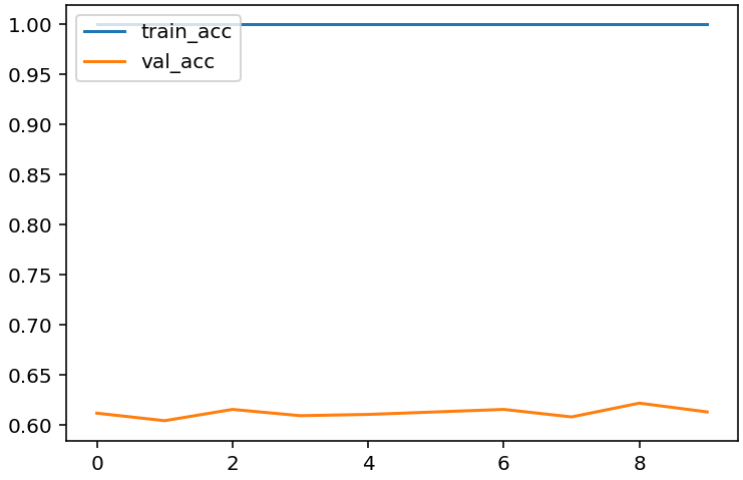


```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=80, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1800, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.6575
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         5
          1       0.00      0.00      0.00        29
          2       0.70      0.80      0.75       334
          3       0.66      0.64      0.65       327
          4       0.52      0.53      0.52        94
          5       0.00      0.00      0.00        11

avg / total       0.62      0.66      0.64       800

The accuarcy for cross_validation
```

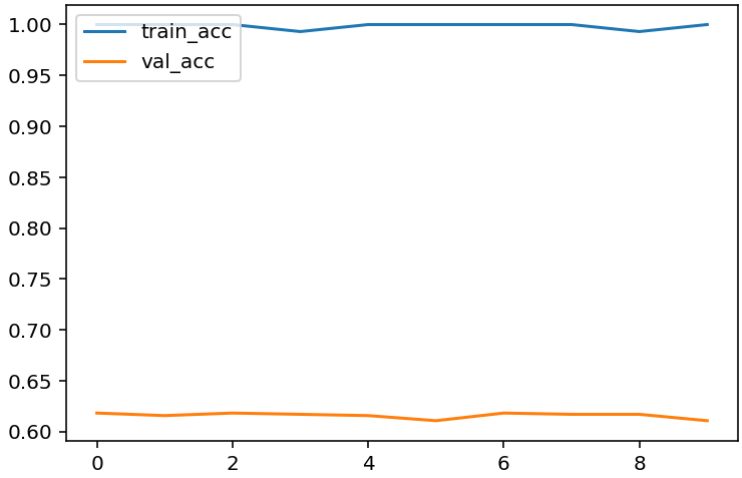


```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=90, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1400, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.66875
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         5
          1       0.00      0.00      0.00        29
          2       0.71      0.81      0.76       334
          3       0.67      0.65      0.66       327
          4       0.54      0.54      0.54        94
          5       0.00      0.00      0.00        11

avg / total       0.63      0.67      0.65       800

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.66375
```

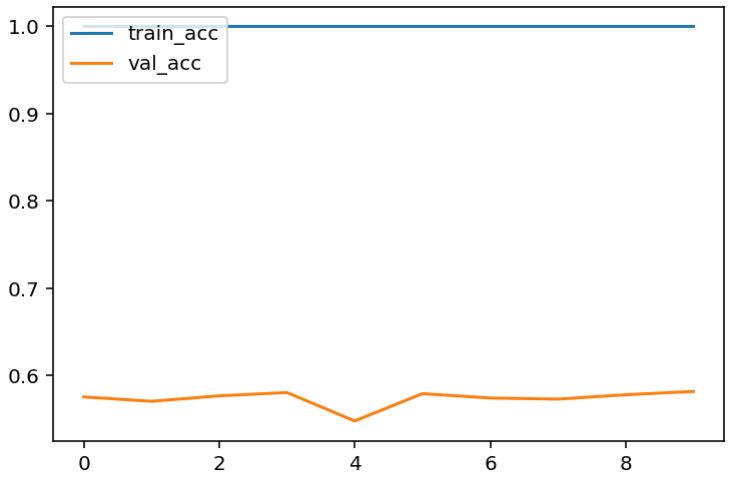
In [14]:
```
#partition2 classifier3
gb_classifier(train32_X, train32_Y, test32_X, test32_Y)
```

```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.15, loss='deviance', max_depth=30,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1400,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.60625
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         5
          1       0.11      0.03      0.05        29
          2       0.67      0.74      0.70       334
          3       0.60      0.58      0.59       327
          4       0.45      0.51      0.48        94
          5       0.00      0.00      0.00        11

avg / total       0.58      0.61      0.59       800

The accuarcy for cross_validation
```

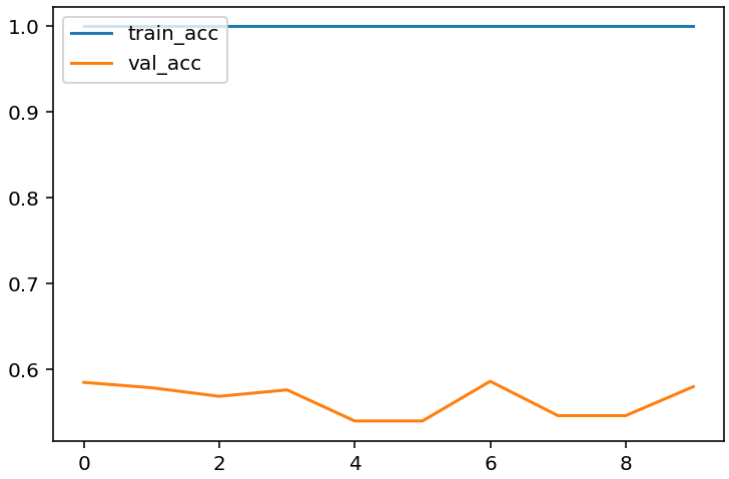


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.15, loss='deviance', max_depth=80,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1200,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.61125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         5
          1       0.20      0.07      0.10        29
          2       0.68      0.74      0.71       334
          3       0.60      0.58      0.59       327
          4       0.45      0.51      0.48        94
          5       0.00      0.00      0.00        11

avg / total       0.59      0.61      0.60       800

The accuarcy for cross_validation
```

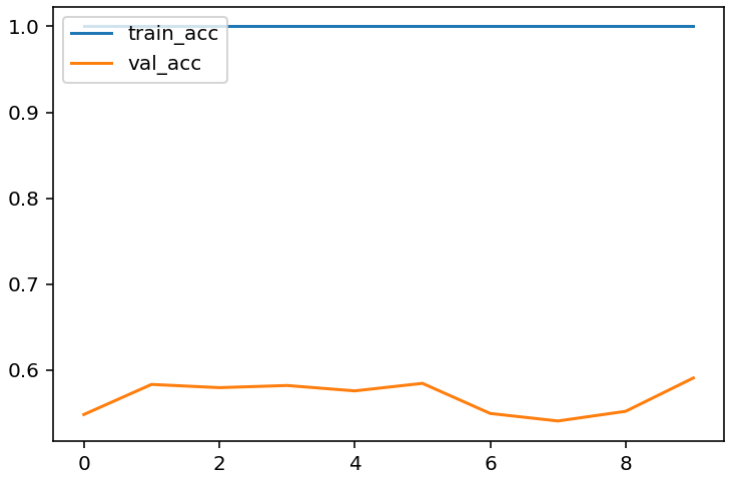


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.15, loss='deviance', max_depth=90,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1400,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.61
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         5
          1       0.10      0.03      0.05        29
          2       0.67      0.74      0.71       334
          3       0.61      0.58      0.59       327
          4       0.46      0.53      0.49        94
          5       0.00      0.00      0.00        11

avg / total       0.59      0.61      0.60       800

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.6091666666666665
```

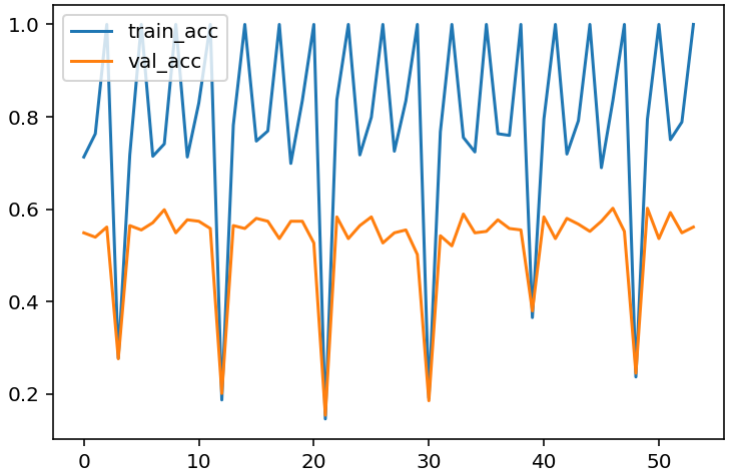
In [13]:
```
#partition3 classifier1
NN_MLP(train33_X, train33_Y, test33_X, test33_Y)
```

```
best estimator is:  MLPClassifier(activation='relu', alpha=1e-06, batch
_size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='adam', tol=0.0001, validation_fraction=0.
1,
       verbose=False, warm_start=False)
test accuracy is: 0.53828125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         8
          1       0.12      0.07      0.09        44
          2       0.61      0.69      0.65       547
          3       0.52      0.44      0.47       514
          4       0.42      0.55      0.47       153
          5       0.00      0.00      0.00        14

avg / total       0.52      0.54      0.53      1280

The accuarcy for cross_validation
```

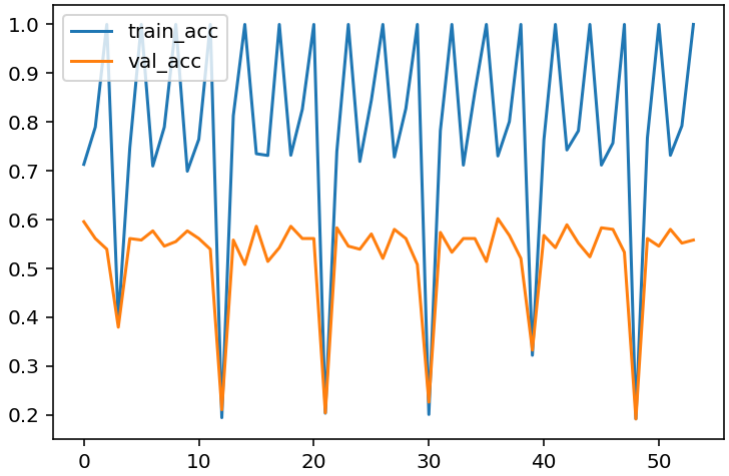


```
best estimator is:  MLPClassifier(activation='relu', alpha=1e-05, batch
_size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.58203125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         8
          1       0.29      0.05      0.08        44
          2       0.64      0.72      0.68       547
          3       0.55      0.53      0.54       514
          4       0.47      0.50      0.48       153
          5       0.00      0.00      0.00        14

avg / total       0.56      0.58      0.57      1280

The accuarcy for cross_validation
```

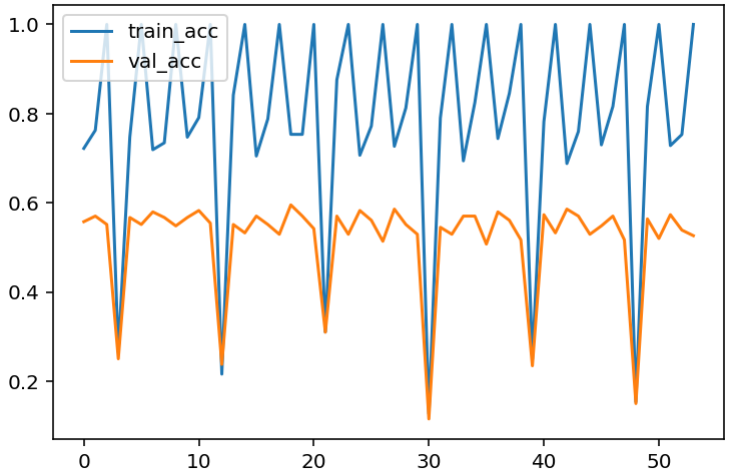


```
best estimator is:  MLPClassifier(activation='relu', alpha=0.001, batch
_size='auto', beta_1=0.9,
       beta_2=0.999, early_stopping=False, epsilon=1e-08,
       hidden_layer_sizes=(100, 100), learning_rate='constant',
       learning_rate_init=0.001, max_iter=1000, momentum=0.9,
       nesterovs_momentum=True, power_t=0.5, random_state=None,
       shuffle=True, solver='sgd', tol=0.0001, validation_fraction=0.1,
       verbose=False, warm_start=False)
test accuracy is: 0.578125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         8
          1       0.30      0.07      0.11        44
          2       0.65      0.71      0.68       547
          3       0.54      0.52      0.53       514
          4       0.47      0.53      0.50       153
          5       0.00      0.00      0.00        14

avg / total       0.56      0.58      0.57      1280

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.5661458333333333
```

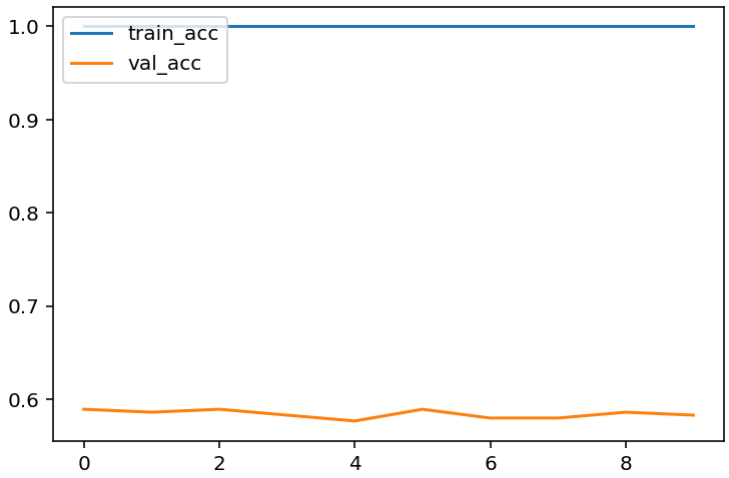
In [11]:
```
#partition3 classifier2
random_forest_classifier(train33_X, train33_Y, test33_X, test33_Y)
```

```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=100, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1400, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.5953125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         8
          1       0.00      0.00      0.00        44
          2       0.66      0.73      0.70       547
          3       0.56      0.56      0.56       514
          4       0.46      0.47      0.46       153
          5       0.00      0.00      0.00        14

avg / total       0.56      0.60      0.58      1280

The accuarcy for cross_validation
```

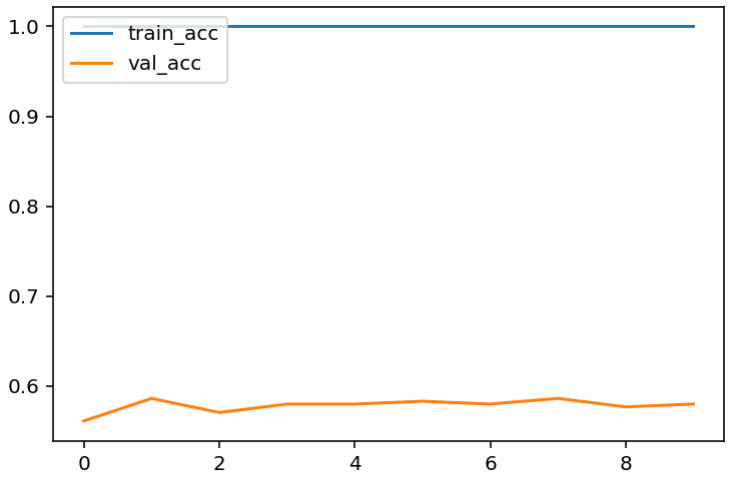


```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=50, max_features='sqrt', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1000, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.5984375
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         8
          1       0.00      0.00      0.00        44
          2       0.66      0.72      0.69       547
          3       0.57      0.57      0.57       514
          4       0.47      0.49      0.48       153
          5       0.00      0.00      0.00        14

avg / total       0.57      0.60      0.58      1280

The accuarcy for cross_validation
```

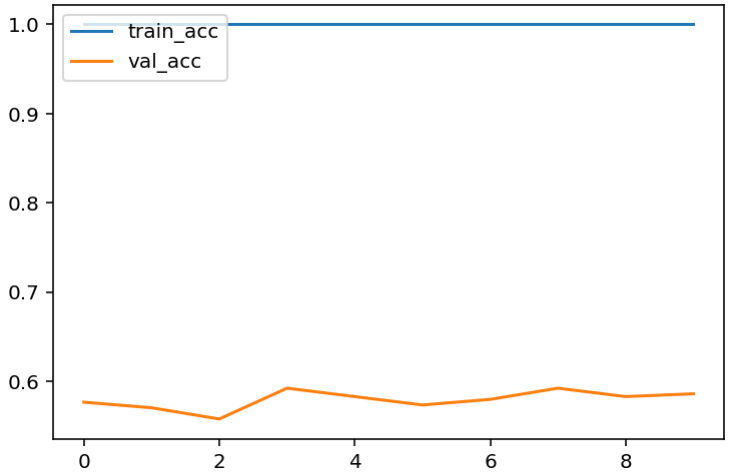


```
best estimator is:  RandomForestClassifier(bootstrap=True, class_weight
=None, criterion='gini',
            max_depth=10, max_features='auto', max_leaf_nodes=None,
            min_impurity_decrease=0.0, min_impurity_split=None,
            min_samples_leaf=1, min_samples_split=2,
            min_weight_fraction_leaf=0.0, n_estimators=1000, n_jobs=1,
            oob_score=False, random_state=None, verbose=0,
            warm_start=False)
test accuracy is: 0.6078125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         8
          1       0.00      0.00      0.00        44
          2       0.67      0.73      0.70       547
          3       0.58      0.59      0.58       514
          4       0.48      0.48      0.48       153
          5       0.00      0.00      0.00        14

avg / total       0.58      0.61      0.59      1280

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.6005208333333334
```

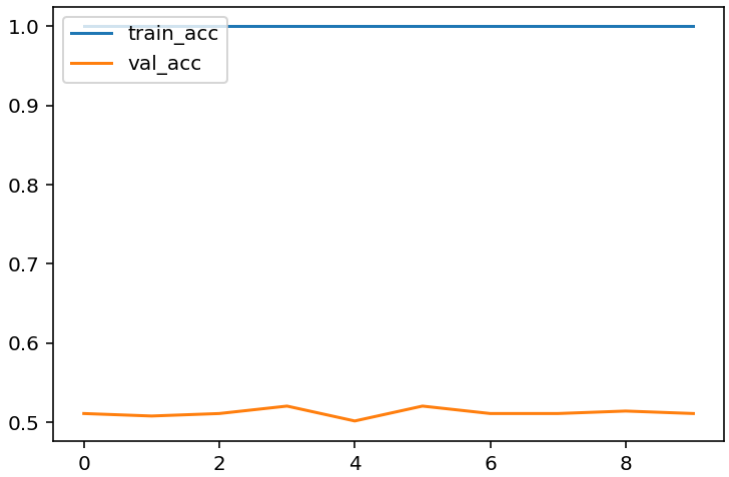
In [12]:
```
#partition3 classifier3
gb_classifier(train33_X, train33_Y, test33_X, test33_Y)
```

```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.05, loss='deviance', max_depth=60,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=600,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.54296875
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         8
          1       0.10      0.07      0.08        44
          2       0.63      0.64      0.64       547
          3       0.51      0.54      0.52       514
          4       0.42      0.44      0.43       153
          5       0.00      0.00      0.00        14

avg / total       0.53      0.54      0.54      1280

The accuarcy for cross_validation
```

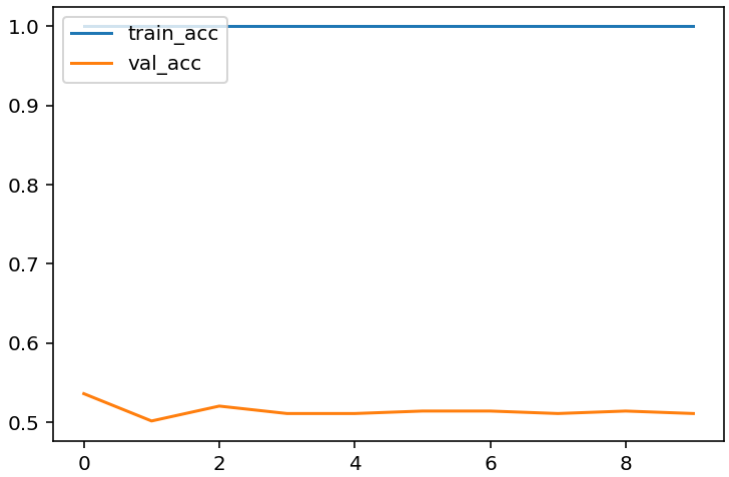


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.005, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1600,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.54453125
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         8
          1       0.08      0.05      0.06        44
          2       0.62      0.67      0.65       547
          3       0.52      0.51      0.52       514
          4       0.40      0.42      0.41       153
          5       0.00      0.00      0.00        14

avg / total       0.53      0.54      0.54      1280

The accuarcy for cross_validation
```

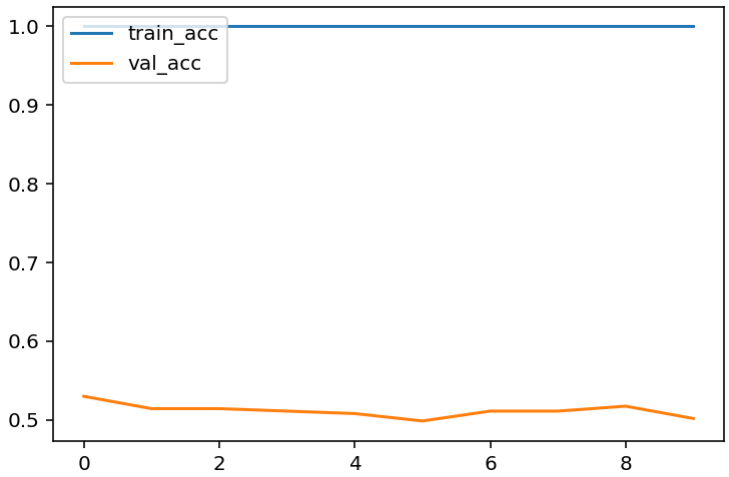


```
best estimator is:  GradientBoostingClassifier(criterion='friedman_ms
e', init=None,
              learning_rate=0.005, loss='deviance', max_depth=10,
              max_features=None, max_leaf_nodes=None,
              min_impurity_decrease=0.0, min_impurity_split=None,
              min_samples_leaf=1, min_samples_split=2,
              min_weight_fraction_leaf=0.0, n_estimators=1200,
              presort='auto', random_state=None, subsample=1.0, verbose
=0,
              warm_start=False)
test accuracy is: 0.5421875
             precision    recall  f1-score   support

          0       0.00      0.00      0.00         8
          1       0.08      0.05      0.06        44
          2       0.62      0.67      0.64       547
          3       0.52      0.51      0.52       514
          4       0.40      0.42      0.41       153
          5       0.00      0.00      0.00        14

avg / total       0.53      0.54      0.53      1280

The accuarcy for cross_validation
```




```
average of test_accuracy is: 0.5432291666666667
```