Class Ex.1 by Matan Porat and Rotem Tsalisher

Guidlines:

- All functions are implemented in the notebook
- Names of functions are determined by the question number
- All functions are called in the main function (last function in the notebook)

In [1]:
```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.discriminant_analysis import QuadraticDiscriminantAnalysis as QDA
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis as LDA
from sklearn.naive_bayes import GaussianNB as GNB
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
```

Part 1

In [2]:
```
def q2():

    X = pd.read_csv('features.csv')
    Y = pd.read_csv('labels.csv')

    return X.values,Y.values
```

In [3]:
```
def q3(X):
    plt.imshow(X[:50],cmap='gray');
    return
```

In [4]:
```
def q4(X,Y):
    xtrain,xtest,ytrain,ytest = train_test_split(X,Y,test_size = 0.2, random_state
    return xtrain,xtest,ytrain[:,0],ytest[:,0];
```

Part 2

In [5]:
```
def q5(x,y,type_):
    if(type_ == "QDA"):
        clf = QDA() # builds an object for results

    elif(type_ == "LDA"):
        clf = LDA();

    elif(type_ == "GNB"):
        clf = GNB();

    clf.fit(x,y) # study data and it's parameters
    return clf.predict(x) # classify
```

In [6]:
```
def q6(y,y_,string,type_):
    cm = confusion_matrix(y,y_)
    print("\n================================");
    print("%s Type Classifier:"%(type_));
```

```
    print("================================");
    print("Q6: %s Set classification accuracy = "%(string), np.diagonal(cm).sum()/s
    print("\n%s Set Confusion Matrix\n\n"%(string),cm)
    print("\n%s Set full Classification Report:\n\n"%(string),classification_report
    return;
```

In [7]:
```
def main():

    # for part 1:

    # q2
    X,Y = q2();

    # q3
    q3(X)

    # q4
    xtrain,xtest,ytrain,ytest = q4(X,Y);

    # for part 2:

    # q5
    type_ = "QDA";
    ytrain_ = q5(xtrain,ytrain,type_);
    ytest_ = q5(xtest,ytest,type_);

    # q6
    q6(ytrain,ytrain_,"Train",type_);
    q6(ytest,ytest_,"Test",type_);

    # q7
    type_ = "LDA";
    ytrain_ = q5(xtrain,ytrain,type_);
    ytest_ = q5(xtest,ytest,type_);
    q6(ytrain,ytrain_,"Train",type_);
    q6(ytest,ytest_,"Test",type_);

    # q8
    type_ = "GNB";
    ytrain_ = q5(xtrain,ytrain,type_);
    ytest_ = q5(xtest,ytest,type_);
    q6(ytrain,ytrain_,"Train",type_);
    q6(ytest,ytest_,"Test",type_);

    return
```

In [8]:
```
main()
```

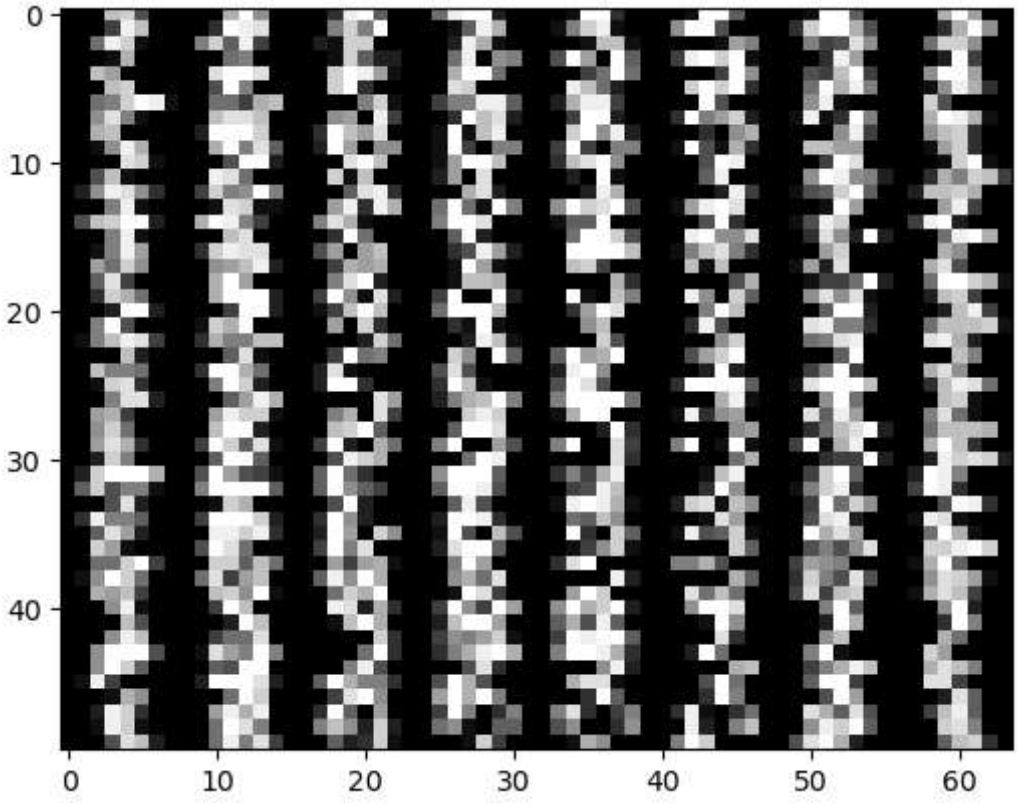
```
C:\Users\420\anaconda3\Lib\site-packages\sklearn\discriminant_analysis.py:935: Use
rWarning: Variables are collinear
  warnings.warn("Variables are collinear")
C:\Users\420\anaconda3\Lib\site-packages\sklearn\discriminant_analysis.py:935: Use
rWarning: Variables are collinear
  warnings.warn("Variables are collinear")
```

```
================================
QDA Type Classifier:
================================
Q6: Train Set classification accuracy =  0.9025069637883009

Train Set Confusion Matrix

 [[137   0   0   0   0   0   0   0   0   0]
 [  0 132   0   0   0   0   0   3  17   0]
 [  0   1 138   0   0   0   0   0   6   0]
 [  0   0   0 115   0   3   0   1  33   0]
 [  1   0   0   0 115   1   0  24   6   0]
 [  0   0   0   0   0 143   0   0   1   0]
 [  0   0   0   0   0   6 120   0  10   0]
 [  0   0   0   0   0   0   0 142   0   0]
 [  0   0   0   0   0   0   0   0 134   0]
 [  0   1   0   0   0   8   0   7  11 120]]

Train Set full Classification Report:

              precision    recall  f1-score   support

           0       0.99      1.00      1.00       137
           1       0.99      0.87      0.92       152
           2       1.00      0.95      0.98       145
           3       1.00      0.76      0.86       152
           4       1.00      0.78      0.88       147
           5       0.89      0.99      0.94       144
           6       1.00      0.88      0.94       136
           7       0.80      1.00      0.89       142
           8       0.61      1.00      0.76       134
           9       1.00      0.82      0.90       147

    accuracy                           0.90      1436
   macro avg       0.93      0.91      0.91      1436
weighted avg       0.93      0.90      0.91      1436


================================
QDA Type Classifier:
================================
Q6: Test Set classification accuracy =  1.0

Test Set Confusion Matrix

 [[40  0  0  0  0  0  0  0  0  0]
 [ 0 30  0  0  0  0  0  0  0  0]
 [ 0  0 32  0  0  0  0  0  0  0]
 [ 0  0  0 31  0  0  0  0  0  0]
 [ 0  0  0  0 34  0  0  0  0  0]
 [ 0  0  0  0  0 38  0  0  0  0]
 [ 0  0  0  0  0  0 45  0  0  0]
 [ 0  0  0  0  0  0  0 37  0  0]
 [ 0  0  0  0  0  0  0  0 40  0]
 [ 0  0  0  0  0  0  0  0  0 33]]

Test Set full Classification Report:

              precision    recall  f1-score   support

           0       1.00      1.00      1.00        40
           1       1.00      1.00      1.00        30
           2       1.00      1.00      1.00        32
           3       1.00      1.00      1.00        31
```

```
           4       1.00      1.00      1.00        34
           5       1.00      1.00      1.00        38
           6       1.00      1.00      1.00        45
           7       1.00      1.00      1.00        37
           8       1.00      1.00      1.00        40
           9       1.00      1.00      1.00        33

    accuracy                           1.00       360
   macro avg       1.00      1.00      1.00       360
weighted avg       1.00      1.00      1.00       360


==================================
LDA Type Classifier:
==================================
Q6: Train Set classification accuracy =  0.9686629526462396

Train Set Confusion Matrix

 [[136   0   0   0   0   0   1   0   0   0]
 [  0 146   0   0   0   0   0   0   2   4]
 [  0   0 142   3   0   0   0   0   0   0]
 [  0   0   1 144   0   2   0   0   4   1]
 [  0   1   0   0 143   0   0   0   1   2]
 [  0   0   0   0   0 141   0   0   0   3]
 [  0   1   0   0   0   0 134   0   0   1]
 [  0   0   0   0   0   0   0 140   0   2]
 [  0   7   0   0   0   0   0   1 124   2]
 [  0   0   0   1   0   3   0   0   2 141]]

Train Set full Classification Report:

              precision    recall  f1-score   support

           0       1.00      0.99      1.00       137
           1       0.94      0.96      0.95       152
           2       0.99      0.98      0.99       145
           3       0.97      0.95      0.96       152
           4       1.00      0.97      0.99       147
           5       0.97      0.98      0.97       144
           6       0.99      0.99      0.99       136
           7       0.99      0.99      0.99       142
           8       0.93      0.93      0.93       134
           9       0.90      0.96      0.93       147

    accuracy                           0.97      1436
   macro avg       0.97      0.97      0.97      1436
weighted avg       0.97      0.97      0.97      1436


==================================
LDA Type Classifier:
==================================
Q6: Test Set classification accuracy =  0.9833333333333333

Test Set Confusion Matrix

 [[40  0  0  0  0  0  0  0  0  0]
 [ 0 29  0  0  0  0  0  0  0  1]
 [ 0  0 32  0  0  0  0  0  0  0]
 [ 0  0  0 30  0  0  0  0  0  1]
 [ 0  0  0  0 33  0  0  1  0  0]
 [ 0  0  0  0  0 37  0  0  0  1]
 [ 0  0  0  0  0  0 45  0  0  0]
```

```
 [ 0  0  0  0  0  0  0 37  0  0]
 [ 0  1  0  0  0  0  0  0 39  0]
 [ 0  0  0  0  0  0  0  0  1 32]]

Test Set full Classification Report:

              precision    recall  f1-score   support

           0       1.00      1.00      1.00        40
           1       0.97      0.97      0.97        30
           2       1.00      1.00      1.00        32
           3       1.00      0.97      0.98        31
           4       1.00      0.97      0.99        34
           5       1.00      0.97      0.99        38
           6       1.00      1.00      1.00        45
           7       0.97      1.00      0.99        37
           8       0.97      0.97      0.97        40
           9       0.91      0.97      0.94        33

    accuracy                           0.98       360
   macro avg       0.98      0.98      0.98       360
weighted avg       0.98      0.98      0.98       360


==================================
GNB Type Classifier:
==================================
Q6: Train Set classification accuracy =  0.8391364902506964

Train Set Confusion Matrix

 [[135   0   0   0   0   0   0   1   1   0]
 [  0 125   0   0   0   0   1   0  22   4]
 [  0   8  82   0   0   0   0   0  55   0]
 [  0   0   1  96   0   7   0   6  40   2]
 [  1   1   0   0 124   2   2  13   4   0]
 [  0   0   0   1   0 136   1   3   3   0]
 [  0   0   0   0   1   1 133   0   1   0]
 [  0   0   0   0   1   0   0 141   0   0]
 [  0   3   0   0   0   1   0   1 129   0]
 [  1   5   0   4   0   2   0  13  18 104]]

Train Set full Classification Report:

              precision    recall  f1-score   support

           0       0.99      0.99      0.99       137
           1       0.88      0.82      0.85       152
           2       0.99      0.57      0.72       145
           3       0.95      0.63      0.76       152
           4       0.98      0.84      0.91       147
           5       0.91      0.94      0.93       144
           6       0.97      0.98      0.97       136
           7       0.79      0.99      0.88       142
           8       0.47      0.96      0.63       134
           9       0.95      0.71      0.81       147

    accuracy                           0.84      1436
   macro avg       0.89      0.84      0.84      1436
weighted avg       0.89      0.84      0.84      1436


==================================
GNB Type Classifier:
```

```
==================================
Q6: Test Set classification accuracy =  0.9

Test Set Confusion Matrix

 [[40  0  0  0  0  0  0  0  0  0]
 [ 0 28  0  0  0  0  1  1  0  0]
 [ 0  3 29  0  0  0  0  0  0  0]
 [ 0  0  0 29  0  0  0  1  1  0]
 [ 0  0  0  0 31  0  0  3  0  0]
 [ 0  0  0  0  0 35  1  1  0  1]
 [ 0  0  0  0  0  0 45  0  0  0]
 [ 0  0  0  0  0  0  0 37  0  0]
 [ 0  5  1  1  0  0  0  2 31  0]
 [ 0  2  1  5  0  1  0  2  3 19]]

Test Set full Classification Report:

              precision    recall  f1-score   support

           0       1.00      1.00      1.00        40
           1       0.74      0.93      0.82        30
           2       0.94      0.91      0.92        32
           3       0.83      0.94      0.88        31
           4       1.00      0.91      0.95        34
           5       0.97      0.92      0.95        38
           6       0.96      1.00      0.98        45
           7       0.79      1.00      0.88        37
           8       0.89      0.78      0.83        40
           9       0.95      0.58      0.72        33

    accuracy                           0.90       360
   macro avg       0.91      0.90      0.89       360
weighted avg       0.91      0.90      0.90       360
```



In [ ]:

In [ ]:

In [ ]: