Homework #3

Due on November 30, 2019 at 11:55pm

Professor Hoda Mohammadzade

Amirhossein Yousefi

97206984

Problem 6

Section a:
In these parts of problem ,we evaluate some classification methods on *digits* which available in **scikit** library.In this section digit data is shown in below figures.

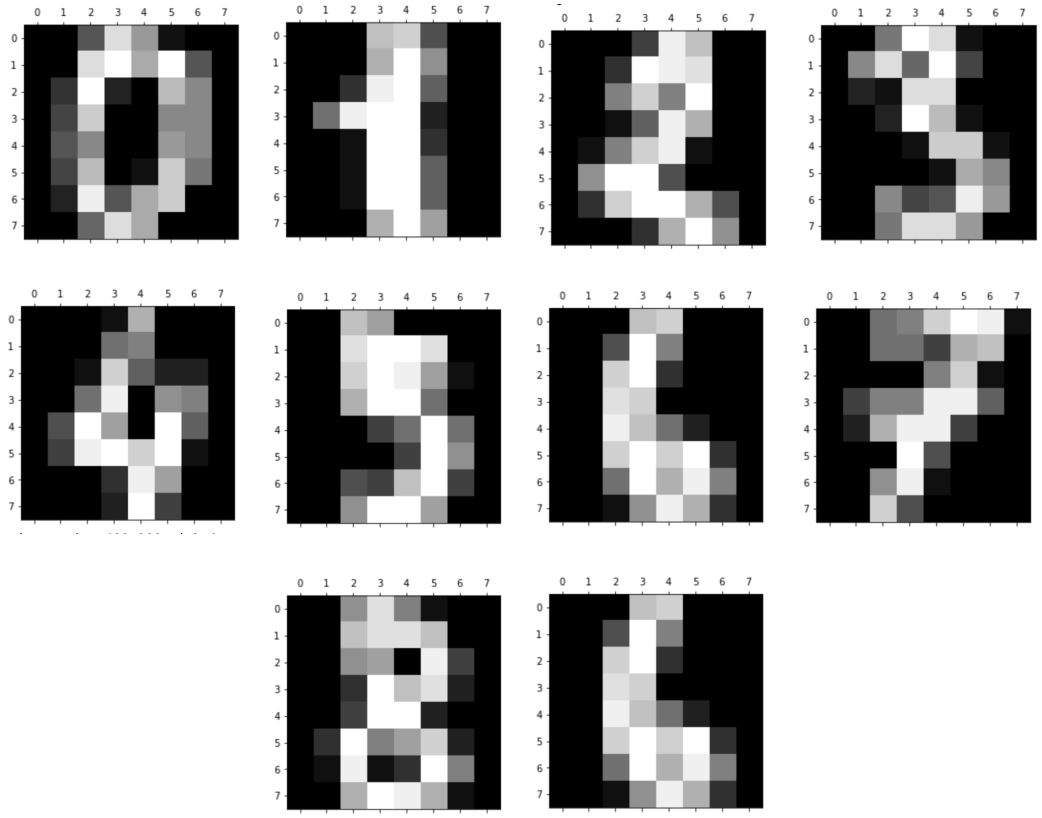


Figure 1: Digits

Section b:
Results for various classification methods is shown bellow.As we can see logistic regression did better and achieved the best accuracy among others.

```
accuracy for QDA is 0.8688888888888889
[[37  1  0  0  0  0  0  0  0  0]
 [ 0 35  0  0  0  0  0  0  2  0]
 [ 0  2 43  0  1  0  0  0  2  0]
 [ 0  0  0 43  0  0  0  2  2  0]
 [ 0  1  0  0 33  0  0  5  0  0]
 [ 0  0  0 10  1 32  0  2  7  1]
 [ 0  2  0  0  0  0 42  0  2  0]
 [ 0  0  0  0  0  0  0 48  0  0]
 [ 0  3  0  1  0  0  0  1 48  0]
 [ 0  0  0  3  0  1  0  5  2 30]]
              precision    recall  f1-score   support

           0       1.00      0.97      0.99        38
           1       0.80      0.95      0.86        37
           2       1.00      0.90      0.95        48
           3       0.75      0.91      0.83        47
           4       0.94      0.85      0.89        39
           5       0.97      0.60      0.74        53
           6       1.00      0.91      0.95        46
           7       0.76      1.00      0.86        48
           8       0.74      0.91      0.81        53
           9       0.97      0.73      0.83        41

    accuracy                           0.87       450
   macro avg       0.89      0.87      0.87       450
weighted avg       0.89      0.87      0.87       450
```

Figure 2: QDA results

```
accuracy for LDA is 0.94
[[38  0  0  0  0  0  0  0  0  0]
 [ 0 34  1  0  0  0  0  0  1  1]
 [ 0  0 48  0  0  0  0  0  0  0]
 [ 0  0  0 43  0  1  0  0  2  1]
 [ 0  1  0  0 37  0  0  1  0  0]
 [ 0  0  0  0  0 53  0  0  0  0]
 [ 0  1  0  0  0  0 44  0  0  1]
 [ 0  0  0  0  0  0  0 47  0  1]
 [ 0  7  0  0  0  0  0  0 44  2]
 [ 0  0  0  0  0  1  0  2  3 35]]
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        38
           1       0.79      0.92      0.85        37
           2       0.98      1.00      0.99        48
           3       1.00      0.91      0.96        47
           4       1.00      0.95      0.97        39
           5       0.96      1.00      0.98        53
           6       1.00      0.96      0.98        46
           7       0.94      0.98      0.96        48
           8       0.88      0.83      0.85        53
           9       0.85      0.85      0.85        41

    accuracy                           0.94       450
   macro avg       0.94      0.94      0.94       450
weighted avg       0.94      0.94      0.94       450
```

Figure 3: LDA results

Problem 6 continued on next page...

```
accuracy for Logistic Regression is 0.96
[[38  0  0  0  0  0  0  0  0  0]
 [ 0 36  0  0  1  0  0  0  0  0]
 [ 0  0 48  0  0  0  0  0  0  0]
 [ 0  0  0 47  0  0  0  0  0  0]
 [ 0  1  0  0 38  0  0  0  0  0]
 [ 0  0  0  0  0 53  0  0  0  0]
 [ 0  1  0  0  0  1 43  0  1  0]
 [ 0  0  0  0  1  1  0 46  0  0]
 [ 0  3  1  0  0  0  0  0 47  2]
 [ 0  0  0  1  1  1  0  1  1 36]]
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        38
           1       0.88      0.97      0.92        37
           2       0.98      1.00      0.99        48
           3       0.98      1.00      0.99        47
           4       0.93      0.97      0.95        39
           5       0.95      1.00      0.97        53
           6       1.00      0.93      0.97        46
           7       0.98      0.96      0.97        48
           8       0.96      0.89      0.92        53
           9       0.95      0.88      0.91        41

    accuracy                           0.96       450
   macro avg       0.96      0.96      0.96       450
weighted avg       0.96      0.96      0.96       450
```

Figure 4: Logistic regression

After logistic regression ,LDA did best on digits data and QDA has the lowest accuracy.

Section c:

Confusion matrix is plotted bellow and accuracy for every digits also has shown.

	precision	recall	f1-score	support
0	1.00	0.97	0.99	38
1	0.80	0.95	0.86	37
2	1.00	0.90	0.95	48
3	0.75	0.91	0.83	47
4	0.94	0.85	0.89	39
5	0.97	0.60	0.74	53
6	1.00	0.91	0.95	46
7	0.76	1.00	0.86	48
8	0.74	0.91	0.81	53
9	0.97	0.73	0.83	41
accuracy			0.87	450
macro avg	0.89	0.87	0.87	450
weighted avg	0.89	0.87	0.87	450

(a) QDA

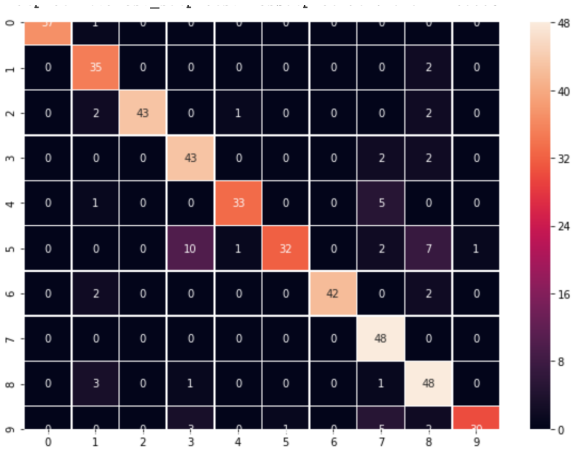
	precision	recall	f1-score	support
0	1.00	1.00	1.00	38
1	0.79	0.92	0.85	37
2	0.98	1.00	0.99	48
3	1.00	0.91	0.96	47
4	1.00	0.95	0.97	39
5	0.96	1.00	0.98	53
6	1.00	0.96	0.98	46
7	0.94	0.98	0.96	48
8	0.88	0.83	0.85	53
9	0.85	0.85	0.85	41
accuracy			0.94	450
macro avg	0.94	0.94	0.94	450
weighted avg	0.94	0.94	0.94	450

(b) LDA

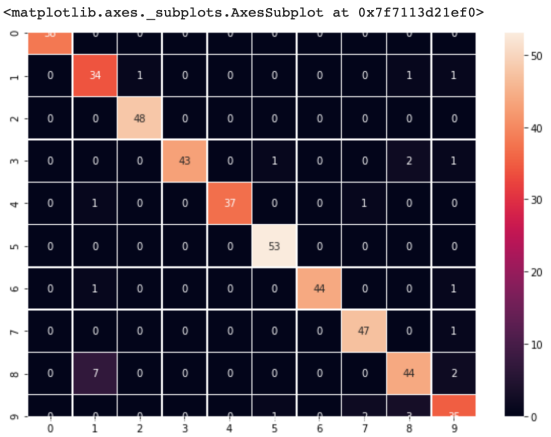
	precision	recall	f1-score	support
0	1.00	1.00	1.00	38
1	0.88	0.97	0.92	37
2	0.98	1.00	0.99	48
3	0.98	1.00	0.99	47
4	0.93	0.97	0.95	39
5	0.95	1.00	0.97	53
6	1.00	0.93	0.97	46
7	0.98	0.96	0.97	48
8	0.96	0.89	0.92	53
9	0.95	0.88	0.91	41
accuracy			0.96	450
macro avg	0.96	0.96	0.96	450
weighted avg	0.96	0.96	0.96	450

(c) Logistic regression

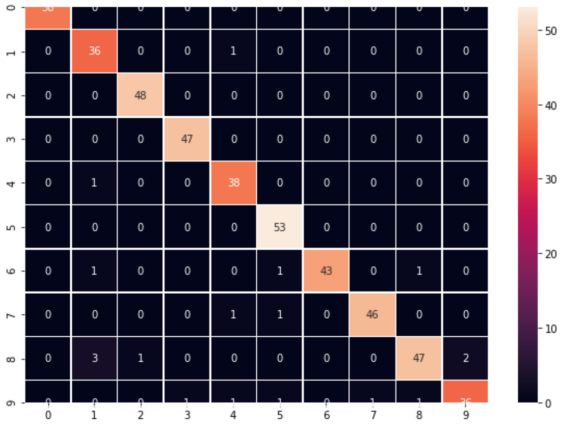
For better illustration notice bellow figures.



(a) QDA



(b) LDA



(c) Logistic regression