Homework4

Point a.
In Figure 1 are shown examples of the digits plotted as images.

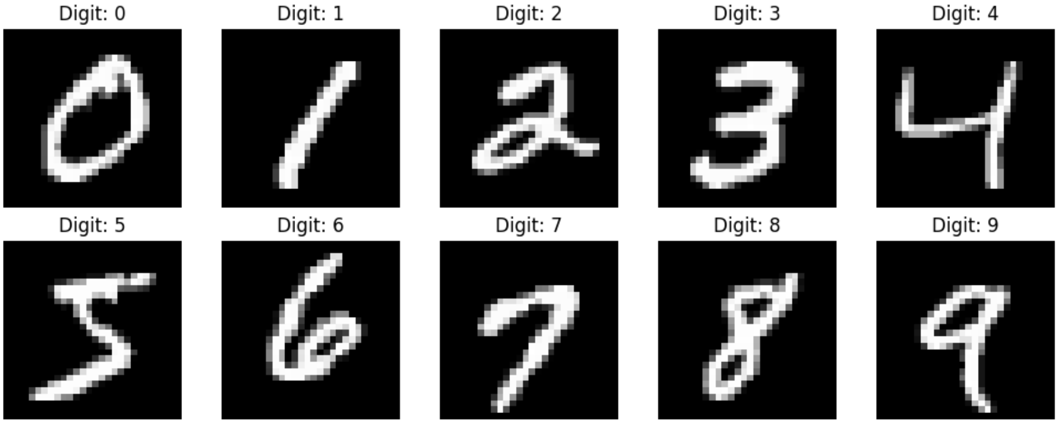

Figure 1

Point c.
Given that the possible outcomes of Y are in [0, 9], if we were to randomly guess a given digit we would get it right once in 10 times. This means that we would get it wrong the remaining 9/10, so on 70000 digits in the dataset we would make, on average, 70000*9/10 = 63000 mistakes.
By looking at the results in Figure 2 we can see that the worst case is when d=10, but we still manage to make roughly half the mistakes we would make if we were to randomly guess the digit.
The value of d for which we get the best predictions is 500 since we get both the lowest MSE and the lowest number of mistakes. Since d is a sort of compressing/expansion factor to reduce/increase the number of elements for each row, of course a higher value for d leads to better performances since we keep more data and the prediction is more accurate. As we reduce the value of d, we get faster (less computationally expensive) results which are worse, though, since we reduced the amount of data.
It's also worth noting that, after a certain value (around 200), an increase of d corresponds to a low improvement in the MSE and the number of mistakes. As we can see from the reported results, the difference of mistakes from d=200 and d=500 is quite low given an increase of d of 150%.

```
d = 10: MSE = 0.765380942035108, #mistakes = 37219
d = 50: MSE = 0.5368059512464681, #mistakes = 16306
d = 100: MSE = 0.46280479610169223, #mistakes = 12236
d = 200: MSE = 0.42098573932184036, #mistakes = 10983
d = 500: MSE = 0.39378594869942474, #mistakes = 10412
```

Figure 2

Point d.

From Figure 3 we can see that we obtain a higher MSE (22% higher) as well as a higher number of mistakes (40% higher) compared to the previous case with d=100. This is because, in this case, we are estimating the best weights W using LMS, so in just 10 epochs we reach just an approximate W. In Figure 4 are reported the MSE values for the 10 epochs: we can see from the downtrend that the algorithm brings an improvement, which is still not as good as the optimal case.

```
LMS configuration: epochs = 10, eta = 0.001
MSE values: [0.8663602407622094, 0.7789582529164806, 0.7186440987957387, 0.6759018458315582, 0.6447746914314189, 0.6214795886435448, 0.6035811517798509
, 0.5894873636485447, 0.5781396462556341, 0.5688209211997391]
d = 100: MSE = 0.561764340973781, #mistakes = 17095
```

Figure 3

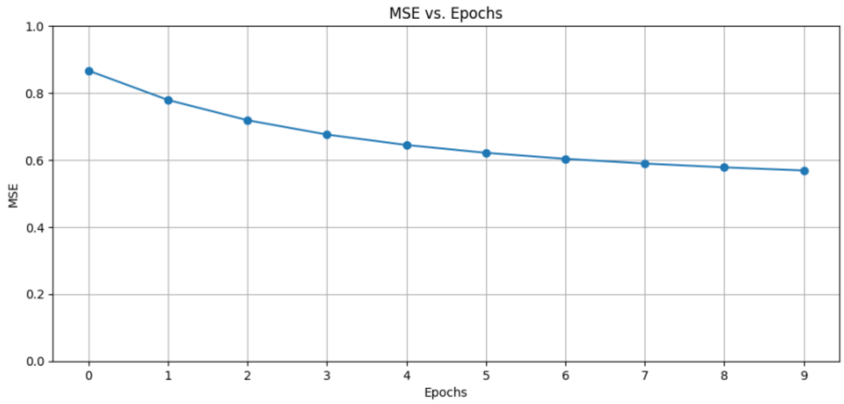

Figure 4

To improve the results, we can try to run the LMS algorithm for more epochs so that it can converge to better values of W. We could also increase the value of eta so that the algorithm converges more quickly.

The results in Figure 5 are run with eta=0.005 and epochs=150. These values are chosen so that the algorithm converges slightly quickly without increasing eta too much and limiting the weights oscillations and with a number of epochs that is a tradeoff between the execution time of the algorithm and the improvement in terms of MSE. With this configuration we can obtain way better results since the MSE and the number of mistakes are very close to the optimal values observed in Point c. In fact, the MSE is 6.5% higher than the optimal case (against the 22% of the previous case) while the number of mistakes is 6.1% higher (against the 40% of the previous case).

In Figure 6 we can see the trend of the MSE values over different epochs.

```
LMS configuration: epochs = 150, eta = 0.005
MSE values: [0.792290426690285, 0.6440913999386567, 0.5992630546797438, 0.5770182076270184, 0.5632780376641277, 0.5538523889040228, 0.5469663735325009,
 0.5417035271174154, 0.5375369907690438, 0.5341435505049945, 0.5313155428689424, 0.5289144382859391, 0.5268447176025232, 0.5250384970185296, 0.52344616
87962272, 0.522030542295913, 0.5207630830654811, 0.5196214408483399, 0.5185877870475811, 0.5176476712136959, 0.5167892171569798, 0.5160025458155838, 0.
5152793525811609, 0.5146125919311481, 0.5139962380567115, 0.5134251003126518, 0.5128946789097034, 0.5124010506313557, 0.5119407772892376, 0.51151083163
71603, 0.511108536857718, 0.5107315167200763, 0.5103776542142946, 0.5100450569816594, 0.5097320282401924, 0.5094370421878209, 0.5091587230806577, 0.508
895827347311, 0.5086472282267128, 0.5084119025159102, 0.5081889190909492, 0.5079774289256164, 0.5077766563814959, 0.5075858915817693, 0.507404483712895
5, 0.5072318351236829, 0.5070673961124466, 0.5069106603098198, 0.5067611605790049, 0.5066184653670331, 0.50648217545008, 0.5063519210243833, 0.50622735
91006837, 0.5061081711662181, 0.5059940610830816, 0.5058847531955805, 0.5057799906232499, 0.5056795337187103, 0.5055831586724142, 0.5054906562485129, 0
.5054018306377167, 0.5053164984151435, 0.505234487592124, 0.505155636752525, 0.505079794264955, 0.5050068175634548, 0.5049365724897977, 0.5048689326916
33, 0.5048037790708301, 0.5047409992776634, 0.5046804872461387, 0.5046221427669074, 0.5045658710942817, 0.5045115825842525, 0.5044591923606255, 0.50440
86200070436, 0.5043597892822906, 0.504312627857149, 0.5042670670708017, 0.5042230417051672, 0.5041804897756207, 0.5041393523367788, 0.5040995733021482,
 0.5040610992763629, 0.5040238793991497, 0.5039878652000025, 0.503953010462716, 0.5039192710990777, 0.5038866050308795, 0.503854972079764, 0.5038243338
642112, 0.5037946537031109, 0.5037658965255293, 0.5037380287861504, 0.5037110183859096, 0.5036848345976412, 0.5036594479961434, 0.5036348303925282, 0.5
036109547724895, 0.5035877952381954, 0.5035653269535877, 0.5035435260928482, 0.5035223697918344, 0.5035018361022138, 0.5034819039482911, 0.503462553086
1042, 0.5034437640648699, 0.5034255181904658, 0.5034077974909597, 0.5033905846839122, 0.5033738631454588, 0.5033576168810163, 0.5033418304974735, 0.503
3264891768638, 0.50331157865139, 0.503297085179668, 0.503282995524179, 0.5032692969299101, 0.5032559771039395, 0.50324302419612 7, 0.5032304267806834, 0
.503218173838635, 0.5032062547411298, 0.5031946592335959, 0.5031833774205277, 0.5031723997510575, 0.5031617170052447, 0.5031513202808652, 0.50314120098
09071, 0.5031313508016406, 0.5031217617211097, 0.5031124259882817, 0.5031033361125956, 0.503094484853987, 0.5030858652133596, 0.503077470423437, 0.5030
692939400873, 0.5030613294339351, 0.5030535707823718, 0.5030460120618926, 0.503038647540804, 0.5030314716721173, 0.5030244790869278, 0.503017664587821,
 0.5030110231428214, 0.5030045498793466, 0.5029982400785746, 0.5029920891699591, 0.5029860927259654, 0.5029802464570657]
d = 100: MSE = 0.4930329691775578, #mistakes = 12987
```

Figure 5

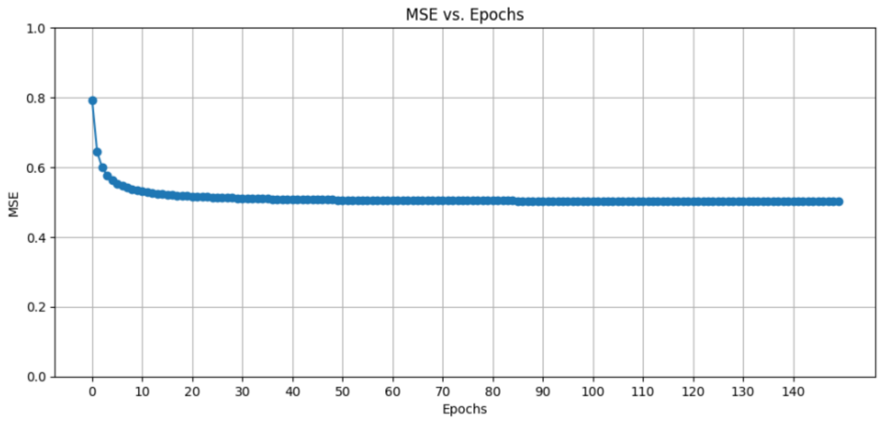
MSE vs. Epochs
MSE
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
Epochs

Figure 6