CSE-6363-007 Assignment 3 (Spring 2024) - Solutions

Mrunmai Nitin Magar (1002092125)

Problem 1: [80 points]

In this problem, we will implement the feedforward and backpropagation process of the neural networks. We will use digital.mat as our experiment data. Then we can train three layer (data, hidden-relu, loss) neural networks and report test accuracy.

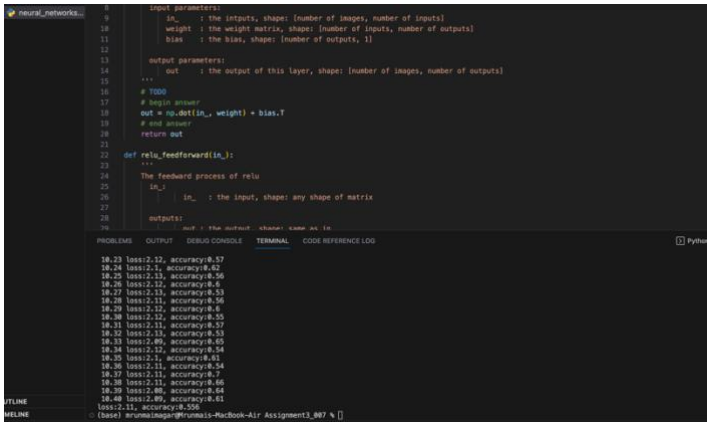
Here is what you need to do:

(i) Finish fullyconnect_feedforward, fullyconnect_backprop, relu_feedforward and relu_backprop part.

(ii) Print loss and accuracy in every training epoch and test.

(iii) *** Do not add or remove libraries in any of the python files. ***

Results:



Complete output:

1. 1 loss:2.3, accuracy:0.12
1. 2 loss:2.3, accuracy:0.12
1. 3 loss:2.32, accuracy:0.07
1. 4 loss:2.33, accuracy:0.09
1. 5 loss:2.29, accuracy:0.15
1. 6 loss:2.31, accuracy:0.11
1. 7 loss:2.31, accuracy:0.1
1. 8 loss:2.3, accuracy:0.13
1. 9 loss:2.3, accuracy:0.06
1.10 loss:2.31, accuracy:0.07
1.11 loss:2.31, accuracy:0.05
1.12 loss:2.32, accuracy:0.08
1.13 loss:2.29, accuracy:0.16
1.14 loss:2.31, accuracy:0.12
1.15 loss:2.31, accuracy:0.15
1.16 loss:2.31, accuracy:0.07
1.17 loss:2.32, accuracy:0.07
1.18 loss:2.33, accuracy:0.07
1.19 loss:2.32, accuracy:0.09
1.20 loss:2.3, accuracy:0.09
1.21 loss:2.29, accuracy:0.14
1.22 loss:2.3, accuracy:0.09
1.23 loss:2.33, accuracy:0.06
1.24 loss:2.33, accuracy:0.07
1.25 loss:2.31, accuracy:0.14
1.26 loss:2.32, accuracy:0.1
1.27 loss:2.29, accuracy:0.13
1.28 loss:2.32, accuracy:0.16
1.29 loss:2.3, accuracy:0.12
1.30 loss:2.31, accuracy:0.13
1.31 loss:2.31, accuracy:0.16
1.32 loss:2.29, accuracy:0.18
1.33 loss:2.3, accuracy:0.15
1.34 loss:2.3, accuracy:0.18
1.35 loss:2.3, accuracy:0.11
1.36 loss:2.32, accuracy:0.13
1.37 loss:2.3, accuracy:0.2
1.38 loss:2.29, accuracy:0.22
1.39 loss:2.3, accuracy:0.18
1.40 loss:2.29, accuracy:0.2
2. 1 loss:2.29, accuracy:0.2
2. 2 loss:2.29, accuracy:0.2
2. 3 loss:2.3, accuracy:0.15
2. 4 loss:2.31, accuracy:0.12
2. 5 loss:2.28, accuracy:0.22
2. 6 loss:2.29, accuracy:0.17
2. 7 loss:2.3, accuracy:0.15
2. 8 loss:2.29, accuracy:0.18
2. 9 loss:2.29, accuracy:0.13
2.10 loss:2.29, accuracy:0.15
2.11 loss:2.3, accuracy:0.12
2.12 loss:2.3, accuracy:0.18
2.13 loss:2.29, accuracy:0.21
2.14 loss:2.3, accuracy:0.18
2.15 loss:2.29, accuracy:0.18
2.16 loss:2.29, accuracy:0.18
2.17 loss:2.3, accuracy:0.15
2.18 loss:2.31, accuracy:0.13
2.19 loss:2.3, accuracy:0.16
2.20 loss:2.29, accuracy:0.21
2.21 loss:2.28, accuracy:0.21
2.22 loss:2.29, accuracy:0.21
2.23 loss:2.3, accuracy:0.15
2.24 loss:2.3, accuracy:0.13

2.25 loss:2.3, accuracy:0.19
2.26 loss:2.3, accuracy:0.14
2.27 loss:2.28, accuracy:0.22
2.28 loss:2.3, accuracy:0.2
2.29 loss:2.29, accuracy:0.19
2.30 loss:2.29, accuracy:0.18
2.31 loss:2.29, accuracy:0.2
2.32 loss:2.28, accuracy:0.18
2.33 loss:2.29, accuracy:0.15
2.34 loss:2.29, accuracy:0.18
2.35 loss:2.29, accuracy:0.13
2.36 loss:2.3, accuracy:0.15
2.37 loss:2.29, accuracy:0.23
2.38 loss:2.28, accuracy:0.27
2.39 loss:2.28, accuracy:0.19
2.40 loss:2.28, accuracy:0.21
3. 1 loss:2.28, accuracy:0.22
3. 2 loss:2.28, accuracy:0.22
3. 3 loss:2.28, accuracy:0.18
3. 4 loss:2.29, accuracy:0.14
3. 5 loss:2.27, accuracy:0.24
3. 6 loss:2.28, accuracy:0.17
3. 7 loss:2.28, accuracy:0.18
3. 8 loss:2.28, accuracy:0.23
3. 9 loss:2.28, accuracy:0.14
3.10 loss:2.28, accuracy:0.2
3.11 loss:2.29, accuracy:0.15
3.12 loss:2.29, accuracy:0.2
3.13 loss:2.28, accuracy:0.23
3.14 loss:2.28, accuracy:0.2
3.15 loss:2.28, accuracy:0.18
3.16 loss:2.28, accuracy:0.22
3.17 loss:2.28, accuracy:0.15
3.18 loss:2.29, accuracy:0.17
3.19 loss:2.29, accuracy:0.17
3.20 loss:2.28, accuracy:0.25
3.21 loss:2.27, accuracy:0.21
3.22 loss:2.28, accuracy:0.24
3.23 loss:2.29, accuracy:0.17
3.24 loss:2.29, accuracy:0.19
3.25 loss:2.28, accuracy:0.2
3.26 loss:2.28, accuracy:0.18
3.27 loss:2.27, accuracy:0.26
3.28 loss:2.28, accuracy:0.22
3.29 loss:2.28, accuracy:0.24
3.30 loss:2.28, accuracy:0.22
3.31 loss:2.28, accuracy:0.21
3.32 loss:2.27, accuracy:0.21
3.33 loss:2.27, accuracy:0.16
3.34 loss:2.28, accuracy:0.28
3.35 loss:2.27, accuracy:0.19
3.36 loss:2.28, accuracy:0.23
3.37 loss:2.27, accuracy:0.35
3.38 loss:2.27, accuracy:0.36
3.39 loss:2.27, accuracy:0.38
3.40 loss:2.27, accuracy:0.34
4. 1 loss:2.27, accuracy:0.31
4. 2 loss:2.27, accuracy:0.36
4. 3 loss:2.27, accuracy:0.34
4. 4 loss:2.28, accuracy:0.26
4. 5 loss:2.26, accuracy:0.46
4. 6 loss:2.27, accuracy:0.42
4. 7 loss:2.27, accuracy:0.34
4. 8 loss:2.26, accuracy:0.42
4. 9 loss:2.27, accuracy:0.37
4.10 loss:2.26, accuracy:0.44
4.11 loss:2.27, accuracy:0.37
4.12 loss:2.27, accuracy:0.39
4.13 loss:2.27, accuracy:0.4
4.14 loss:2.27, accuracy:0.35
4.15 loss:2.27, accuracy:0.41
4.16 loss:2.26, accuracy:0.4
4.17 loss:2.27, accuracy:0.37
4.18 loss:2.27, accuracy:0.34
4.19 loss:2.28, accuracy:0.3
4.20 loss:2.26, accuracy:0.43
4.21 loss:2.26, accuracy:0.45
4.22 loss:2.26, accuracy:0.44
4.23 loss:2.27, accuracy:0.35
4.24 loss:2.27, accuracy:0.4
4.25 loss:2.27, accuracy:0.37
4.26 loss:2.27, accuracy:0.34
4.27 loss:2.26, accuracy:0.41
4.28 loss:2.27, accuracy:0.41
4.29 loss:2.27, accuracy:0.39
4.30 loss:2.27, accuracy:0.4
4.31 loss:2.27, accuracy:0.43
4.32 loss:2.26, accuracy:0.42
4.33 loss:2.26, accuracy:0.44
4.34 loss:2.26, accuracy:0.48
4.35 loss:2.26, accuracy:0.42
4.36 loss:2.27, accuracy:0.39
4.37 loss:2.26, accuracy:0.52
4.38 loss:2.26, accuracy:0.54
4.39 loss:2.25, accuracy:0.57
4.40 loss:2.26, accuracy:0.53
5. 1 loss:2.26, accuracy:0.47
5. 2 loss:2.25, accuracy:0.53
5. 3 loss:2.26, accuracy:0.51
5. 4 loss:2.27, accuracy:0.4
5. 5 loss:2.25, accuracy:0.6
5. 6 loss:2.25, accuracy:0.54
5. 7 loss:2.26, accuracy:0.49
5. 8 loss:2.25, accuracy:0.56
5. 9 loss:2.26, accuracy:0.47
5.10 loss:2.25, accuracy:0.55
5.11 loss:2.26, accuracy:0.48
5.12 loss:2.26, accuracy:0.49
5.13 loss:2.26, accuracy:0.47
5.14 loss:2.26, accuracy:0.42
5.15 loss:2.25, accuracy:0.5
5.16 loss:2.25, accuracy:0.47
5.17 loss:2.25, accuracy:0.52
5.18 loss:2.25, accuracy:0.48
5.19 loss:2.26, accuracy:0.4
5.20 loss:2.25, accuracy:0.5
5.21 loss:2.25, accuracy:0.5
5.22 loss:2.25, accuracy:0.49
5.23 loss:2.25, accuracy:0.49
5.24 loss:2.25, accuracy:0.53
5.25 loss:2.25, accuracy:0.49
5.26 loss:2.25, accuracy:0.49
5.27 loss:2.25, accuracy:0.54
5.28 loss:2.25, accuracy:0.55
5.29 loss:2.25, accuracy:0.53
5.30 loss:2.25, accuracy:0.44
5.31 loss:2.25, accuracy:0.54
5.32 loss:2.25, accuracy:0.49
5.33 loss:2.24, accuracy:0.6
5.34 loss:2.25, accuracy:0.57
5.35 loss:2.24, accuracy:0.54
5.36 loss:2.25, accuracy:0.55
5.37 loss:2.24, accuracy:0.66
5.38 loss:2.24, accuracy:0.63
5.39 loss:2.24, accuracy:0.64
5.40 loss:2.24, accuracy:0.65
6. 1 loss:2.24, accuracy:0.57
6. 2 loss:2.24, accuracy:0.6
6. 3 loss:2.24, accuracy:0.63
6. 4 loss:2.25, accuracy:0.51
6. 5 loss:2.24, accuracy:0.65
6. 6 loss:2.24, accuracy:0.58
6. 7 loss:2.24, accuracy:0.59
6. 8 loss:2.23, accuracy:0.65
6. 9 loss:2.24, accuracy:0.53
6.10 loss:2.23, accuracy:0.59
6.11 loss:2.24, accuracy:0.52
6.12 loss:2.24, accuracy:0.56
6.13 loss:2.24, accuracy:0.51
6.14 loss:2.24, accuracy:0.53
6.15 loss:2.23, accuracy:0.56
6.16 loss:2.23, accuracy:0.62
6.17 loss:2.23, accuracy:0.58
6.18 loss:2.23, accuracy:0.55
6.19 loss:2.25, accuracy:0.49
6.20 loss:2.23, accuracy:0.62
6.21 loss:2.23, accuracy:0.57
6.22 loss:2.23, accuracy:0.54
6.23 loss:2.24, accuracy:0.52
6.24 loss:2.23, accuracy:0.58
6.25 loss:2.24, accuracy:0.57
6.26 loss:2.23, accuracy:0.57
6.27 loss:2.23, accuracy:0.57
6.28 loss:2.23, accuracy:0.56
6.29 loss:2.23, accuracy:0.58
6.30 loss:2.23, accuracy:0.54
6.31 loss:2.23, accuracy:0.6
6.32 loss:2.24, accuracy:0.55
6.33 loss:2.22, accuracy:0.6
6.34 loss:2.23, accuracy:0.61
6.35 loss:2.23, accuracy:0.64
6.36 loss:2.23, accuracy:0.61
6.37 loss:2.23, accuracy:0.68
6.38 loss:2.23, accuracy:0.67
6.39 loss:2.22, accuracy:0.65
6.40 loss:2.22, accuracy:0.67
7. 1 loss:2.23, accuracy:0.58
7. 2 loss:2.22, accuracy:0.6
7. 3 loss:2.22, accuracy:0.64
7. 4 loss:2.23, accuracy:0.55
7. 5 loss:2.22, accuracy:0.65
7. 6 loss:2.21, accuracy:0.64
7. 7 loss:2.22, accuracy:0.65
7. 8 loss:2.21, accuracy:0.64
7. 9 loss:2.22, accuracy:0.55
7.10 loss:2.21, accuracy:0.62
7.11 loss:2.22, accuracy:0.54
7.12 loss:2.22, accuracy:0.61
7.13 loss:2.22, accuracy:0.55
7.14 loss:2.22, accuracy:0.57
7.15 loss:2.21, accuracy:0.65
7.16 loss:2.21, accuracy:0.64
7.17 loss:2.21, accuracy:0.61
7.18 loss:2.21, accuracy:0.6
7.19 loss:2.23, accuracy:0.55
7.20 loss:2.21, accuracy:0.62

7.21 loss:2.21, accuracy:0.61
7.22 loss:2.21, accuracy:0.54
7.23 loss:2.21, accuracy:0.58
7.24 loss:2.21, accuracy:0.61
7.25 loss:2.22, accuracy:0.58
7.26 loss:2.21, accuracy:0.59
7.27 loss:2.22, accuracy:0.56
7.28 loss:2.21, accuracy:0.58
7.29 loss:2.21, accuracy:0.59
7.30 loss:2.21, accuracy:0.55
7.31 loss:2.21, accuracy:0.62
7.32 loss:2.22, accuracy:0.56
7.33 loss:2.2, accuracy:0.62
7.34 loss:2.21, accuracy:0.58
7.35 loss:2.2, accuracy:0.68
7.36 loss:2.21, accuracy:0.57
7.37 loss:2.2, accuracy:0.7
7.38 loss:2.2, accuracy:0.68
7.39 loss:2.19, accuracy:0.64
7.40 loss:2.2, accuracy:0.62
8. 1 loss:2.2, accuracy:0.6
8. 2 loss:2.19, accuracy:0.57
8. 3 loss:2.2, accuracy:0.64
8. 4 loss:2.21, accuracy:0.56
8. 5 loss:2.2, accuracy:0.62
8. 6 loss:2.19, accuracy:0.66
8. 7 loss:2.2, accuracy:0.65
8. 8 loss:2.19, accuracy:0.62
8. 9 loss:2.2, accuracy:0.54
8.10 loss:2.18, accuracy:0.61
8.11 loss:2.2, accuracy:0.57
8.12 loss:2.19, accuracy:0.62
8.13 loss:2.2, accuracy:0.57
8.14 loss:2.19, accuracy:0.58
8.15 loss:2.19, accuracy:0.67
8.16 loss:2.19, accuracy:0.66
8.17 loss:2.18, accuracy:0.58
8.18 loss:2.18, accuracy:0.6
8.19 loss:2.2, accuracy:0.53
8.20 loss:2.19, accuracy:0.62
8.21 loss:2.18, accuracy:0.62
8.22 loss:2.19, accuracy:0.52
8.23 loss:2.19, accuracy:0.57
8.24 loss:2.18, accuracy:0.6
8.25 loss:2.19, accuracy:0.61
8.26 loss:2.19, accuracy:0.6
8.27 loss:2.19, accuracy:0.56
8.28 loss:2.18, accuracy:0.57
8.29 loss:2.19, accuracy:0.6
8.30 loss:2.19, accuracy:0.55
8.31 loss:2.18, accuracy:0.62
8.32 loss:2.19, accuracy:0.55
8.33 loss:2.17, accuracy:0.62
8.34 loss:2.18, accuracy:0.57
8.35 loss:2.17, accuracy:0.65
8.36 loss:2.18, accuracy:0.55
8.37 loss:2.18, accuracy:0.69
8.38 loss:2.18, accuracy:0.66
8.39 loss:2.16, accuracy:0.64
8.40 loss:2.17, accuracy:0.62
9. 1 loss:2.18, accuracy:0.6
9. 2 loss:2.16, accuracy:0.57
9. 3 loss:2.17, accuracy:0.64
9. 4 loss:2.18, accuracy:0.55
9. 5 loss:2.17, accuracy:0.62
9. 6 loss:2.16, accuracy:0.65
9. 7 loss:2.17, accuracy:0.63
9. 8 loss:2.16, accuracy:0.63
9. 9 loss:2.17, accuracy:0.54
9.10 loss:2.14, accuracy:0.61
9.11 loss:2.17, accuracy:0.57
9.12 loss:2.16, accuracy:0.61
9.13 loss:2.17, accuracy:0.59
9.14 loss:2.17, accuracy:0.59
9.15 loss:2.15, accuracy:0.67
9.16 loss:2.16, accuracy:0.63
9.17 loss:2.15, accuracy:0.54
9.18 loss:2.15, accuracy:0.62
9.19 loss:2.18, accuracy:0.53
9.20 loss:2.16, accuracy:0.61
9.21 loss:2.15, accuracy:0.61
9.22 loss:2.15, accuracy:0.52
9.23 loss:2.16, accuracy:0.58
9.24 loss:2.14, accuracy:0.63
9.25 loss:2.16, accuracy:0.57
9.26 loss:2.16, accuracy:0.6
9.27 loss:2.16, accuracy:0.54
9.28 loss:2.15, accuracy:0.57
9.29 loss:2.15, accuracy:0.59
9.30 loss:2.16, accuracy:0.55
9.31 loss:2.15, accuracy:0.61
9.32 loss:2.16, accuracy:0.54
9.33 loss:2.13, accuracy:0.63
9.34 loss:2.15, accuracy:0.55
9.35 loss:2.14, accuracy:0.63
9.36 loss:2.14, accuracy:0.54
9.37 loss:2.14, accuracy:0.69
9.38 loss:2.15, accuracy:0.66
9.39 loss:2.12, accuracy:0.66
9.40 loss:2.13, accuracy:0.62
10. 1 loss:2.15, accuracy:0.6
10. 2 loss:2.13, accuracy:0.56
10. 3 loss:2.14, accuracy:0.63
10. 4 loss:2.15, accuracy:0.53
10. 5 loss:2.14, accuracy:0.62
10. 6 loss:2.12, accuracy:0.63
10. 7 loss:2.14, accuracy:0.62
10. 8 loss:2.12, accuracy:0.63
10. 9 loss:2.13, accuracy:0.55
10.10 loss:2.1, accuracy:0.62
10.11 loss:2.14, accuracy:0.58
10.12 loss:2.12, accuracy:0.6
10.13 loss:2.13, accuracy:0.58
10.14 loss:2.13, accuracy:0.58
10.15 loss:2.12, accuracy:0.65
10.16 loss:2.12, accuracy:0.64
10.17 loss:2.11, accuracy:0.54
10.18 loss:2.12, accuracy:0.6
10.19 loss:2.14, accuracy:0.51
10.20 loss:2.12, accuracy:0.62
10.21 loss:2.11, accuracy:0.6
10.22 loss:2.12, accuracy:0.51
10.23 loss:2.12, accuracy:0.57
10.24 loss:2.1, accuracy:0.62
10.25 loss:2.13, accuracy:0.56
10.26 loss:2.12, accuracy:0.6
10.27 loss:2.13, accuracy:0.53
10.28 loss:2.11, accuracy:0.56
10.29 loss:2.12, accuracy:0.6
10.30 loss:2.12, accuracy:0.55
10.31 loss:2.11, accuracy:0.57
10.32 loss:2.13, accuracy:0.53
10.33 loss:2.09, accuracy:0.65
10.34 loss:2.12, accuracy:0.54
10.35 loss:2.1, accuracy:0.61
10.36 loss:2.11, accuracy:0.54
10.37 loss:2.11, accuracy:0.7
10.38 loss:2.11, accuracy:0.66
10.39 loss:2.08, accuracy:0.64
10.40 loss:2.09, accuracy:0.61
loss:2.11, accuracy:0.556

Problem 2: [20 points]

Let us review the bias-variance decomposition first. The intuition behind it is straight-forward: if the model is too simple, the learnt function is biased and does not fit the data. If the model is too complex then it is very sensitive to small changes in the data. If we were able to sample a dataset D infinite many times, we will learn different g(x) for each time, and get an expected hypothesis $\bar{g}(x)$. So bias means the difference between the truth and what you expect to learn. It measures how well our model can approximate the truth at best. However, it is impossible to sample the training dataset multiple time, so variance means the difference between what you learn from a particular dataset and what you expect to learn. Now please answer the following questions:

True of False:

(i) If a learning algorithm is suffering from high bias, adding more training examples will improve the test error significantly.

(ii) We always prefer models with high variance (over those with high bias) as they will able to better fit the training set.

(iii) A model with more parameters is more prone to overfitting and typically has higher variance.

(iv) Introducing regularization to the model always results in equal or better performance on the training set.

(v) Using a very large value of regularization parameter λ cannot hurt the performance of your hypothesis.

Solution:

(i) False. Adding more training examples typically helps reduce variance, not bias. High bias means the model is too simple and does not fit the data well, irrespective of the number of examples.

(ii) False. Models with high variance are prone to overfitting and may not generalize well to unseen data. It's generally preferred to have a balance between bias and variance.

(iii) True. A model with more parameters has higher capacity and can potentially capture more complex patterns in the data. However, this also increases the risk of overfitting, leading to higher variance.

(iv) False. Introducing regularization may reduce overfitting, but it can also impact performance on the training set. Sometimes, regularization techniques penalize overly complex models, which might reduce their performance on the training set.

(v) False. Using a very large value of the regularization parameter λ can hurt the performance of the hypothesis. It may excessively penalize the model's complexity, leading to underfitting or poor performance on both training and test sets.