**STAT 6910 Assignment 5 Agnibh Dasgupta A02292865**

1.1)

The cost function C (w, λ) is not a function of learning rate η. You can't compute the gradient of η.

λ is part of the cost function. You can indeed compute its gradient and apply SGD on λ. However, when you do that, λ will keep decreasing because its optimal value is obviously −∞.

2.1) Yes, one of the factors responsible for the vanishing gradient problem is the derivative of the activation function being too low resulting in very slow changes leading to convergence. Using one with a relatively higher derivative lik tanh or relu would help with the problem.

2.3) If μ > 1, the network would never converge to the local minima is the gradient would keep oscillating to and fro about the minima but never settle there.

If μ < 0, the network would gain momentum in the opposite direction to the direction it should be taking to reduce the loss, so it would never reach the minima at all.

3.1) **Common parameters**:

num\_epochs = 50

batch\_size = 32

dropout\_rate = 0.2

**SGD** – Learning\_rate = 0.005, Final test accuracy = 83%

**SGD momentum** – learning\_rate = 0.005, Final test accuracy = 93%

**AdaGrad** – learning\_rate = 0.001, Final test accuracy = 57%

**Adam** – learning\_rate = 0.0001, Final test accuracy = 86%

**SGD data:**

Epoch 0. Loss/Training Cost = 3652.83

Epoch 0. Validation Accuracy = 0.16

Epoch 1. Loss/Training Cost = 3649.68

Epoch 2. Loss/Training Cost = 3645.22

Epoch 3. Loss/Training Cost = 3636.38

Epoch 4. Loss/Training Cost = 3603.68

Epoch 5. Loss/Training Cost = 3537.67

Epoch 6. Loss/Training Cost = 3448.42

Epoch 7. Loss/Training Cost = 3355.54

Epoch 8. Loss/Training Cost = 3268.94

Epoch 9. Loss/Training Cost = 3178.76

Epoch 10. Loss/Training Cost = 3090.52

Epoch 10. Validation Accuracy = 0.59

Epoch 11. Loss/Training Cost = 3033.14

Epoch 12. Loss/Training Cost = 2999.73

Epoch 13. Loss/Training Cost = 2974.64

Epoch 14. Loss/Training Cost = 2955.72

Epoch 15. Loss/Training Cost = 2940.24

Epoch 16. Loss/Training Cost = 2922.17

Epoch 17. Loss/Training Cost = 2899.08

Epoch 18. Loss/Training Cost = 2880.26

Epoch 19. Loss/Training Cost = 2865.03

Epoch 20. Loss/Training Cost = 2851.85

Epoch 20. Validation Accuracy = 0.70

Epoch 21. Loss/Training Cost = 2841.65

Epoch 22. Loss/Training Cost = 2832.16

Epoch 23. Loss/Training Cost = 2819.81

Epoch 24. Loss/Training Cost = 2800.46

Epoch 25. Loss/Training Cost = 2780.07

Epoch 26. Loss/Training Cost = 2763.94

Epoch 27. Loss/Training Cost = 2744.32

Epoch 28. Loss/Training Cost = 2731.17

Epoch 29. Loss/Training Cost = 2720.82

Epoch 30. Loss/Training Cost = 2711.88

Epoch 30. Validation Accuracy = 0.79

Epoch 31. Loss/Training Cost = 2704.90

Epoch 32. Loss/Training Cost = 2699.84

Epoch 33. Loss/Training Cost = 2693.54

Epoch 34. Loss/Training Cost = 2689.79

Epoch 35. Loss/Training Cost = 2686.09

Epoch 36. Loss/Training Cost = 2681.89

Epoch 37. Loss/Training Cost = 2679.52

Epoch 38. Loss/Training Cost = 2675.58

Epoch 39. Loss/Training Cost = 2673.09

Epoch 40. Loss/Training Cost = 2668.96

Epoch 40. Validation Accuracy = 0.81

Epoch 41. Loss/Training Cost = 2667.34

Epoch 42. Loss/Training Cost = 2663.98

Epoch 43. Loss/Training Cost = 2662.10

Epoch 44. Loss/Training Cost = 2658.80

Epoch 45. Loss/Training Cost = 2658.24

Epoch 46. Loss/Training Cost = 2656.56

Epoch 47. Loss/Training Cost = 2653.98

Epoch 48. Loss/Training Cost = 2651.63

Epoch 49. Loss/Training Cost = 2650.09

Epoch 50. Loss/Training Cost = 2650.41

Epoch 50. Validation Accuracy = 0.83

**SGD momentum data:**

Epoch 0. Loss/Training Cost = 3481.02

Epoch 0. Validation Accuracy = 0.52

Epoch 1. Loss/Training Cost = 3016.35

Epoch 2. Loss/Training Cost = 2862.90

Epoch 3. Loss/Training Cost = 2748.58

Epoch 4. Loss/Training Cost = 2674.48

Epoch 5. Loss/Training Cost = 2651.70

Epoch 6. Loss/Training Cost = 2632.08

Epoch 7. Loss/Training Cost = 2580.65

Epoch 8. Loss/Training Cost = 2559.33

Epoch 9. Loss/Training Cost = 2543.30

Epoch 10. Loss/Training Cost = 2537.11

Epoch 10. Validation Accuracy = 0.90

Epoch 11. Loss/Training Cost = 2531.41

Epoch 12. Loss/Training Cost = 2523.85

Epoch 13. Loss/Training Cost = 2519.16

Epoch 14. Loss/Training Cost = 2518.00

Epoch 15. Loss/Training Cost = 2515.75

Epoch 16. Loss/Training Cost = 2511.87

Epoch 17. Loss/Training Cost = 2508.83

Epoch 18. Loss/Training Cost = 2507.22

Epoch 19. Loss/Training Cost = 2503.97

Epoch 20. Loss/Training Cost = 2504.13

Epoch 20. Validation Accuracy = 0.92

Epoch 21. Loss/Training Cost = 2501.97

Epoch 22. Loss/Training Cost = 2502.52

Epoch 23. Loss/Training Cost = 2501.43

Epoch 24. Loss/Training Cost = 2499.90

Epoch 25. Loss/Training Cost = 2497.59

Epoch 26. Loss/Training Cost = 2497.87

Epoch 27. Loss/Training Cost = 2495.98

Epoch 28. Loss/Training Cost = 2495.76

Epoch 29. Loss/Training Cost = 2495.78

Epoch 30. Loss/Training Cost = 2494.51

Epoch 30. Validation Accuracy = 0.93

Epoch 31. Loss/Training Cost = 2494.19

Epoch 32. Loss/Training Cost = 2494.08

Epoch 33. Loss/Training Cost = 2493.84

Epoch 34. Loss/Training Cost = 2495.03

Epoch 35. Loss/Training Cost = 2493.18

Epoch 36. Loss/Training Cost = 2494.36

Epoch 37. Loss/Training Cost = 2490.92

Epoch 38. Loss/Training Cost = 2492.08

Epoch 39. Loss/Training Cost = 2493.36

Epoch 40. Loss/Training Cost = 2492.71

Epoch 40. Validation Accuracy = 0.93

Epoch 41. Loss/Training Cost = 2490.87

Epoch 42. Loss/Training Cost = 2491.49

Epoch 43. Loss/Training Cost = 2491.63

Epoch 44. Loss/Training Cost = 2491.18

Epoch 45. Loss/Training Cost = 2490.77

Epoch 46. Loss/Training Cost = 2489.37

Epoch 47. Loss/Training Cost = 2490.68

Epoch 48. Loss/Training Cost = 2491.07

Epoch 49. Loss/Training Cost = 2491.44

Epoch 50. Loss/Training Cost = 2489.23

Epoch 50. Validation Accuracy = 0.93

**SGD AdaGrad data:**

Epoch 0. Loss/Training Cost = 3648.23

Epoch 0. Validation Accuracy = 0.24

Epoch 1. Loss/Training Cost = 3640.65

Epoch 2. Loss/Training Cost = 3632.72

Epoch 3. Loss/Training Cost = 3623.54

Epoch 4. Loss/Training Cost = 3612.84

Epoch 5. Loss/Training Cost = 3600.27

Epoch 6. Loss/Training Cost = 3586.77

Epoch 7. Loss/Training Cost = 3571.60

Epoch 8. Loss/Training Cost = 3554.75

Epoch 9. Loss/Training Cost = 3537.87

Epoch 10. Loss/Training Cost = 3521.33

Epoch 10. Validation Accuracy = 0.41

Epoch 11. Loss/Training Cost = 3504.09

Epoch 12. Loss/Training Cost = 3489.76

Epoch 13. Loss/Training Cost = 3474.81

Epoch 14. Loss/Training Cost = 3461.27

Epoch 15. Loss/Training Cost = 3447.65

Epoch 16. Loss/Training Cost = 3433.27

Epoch 17. Loss/Training Cost = 3422.72

Epoch 18. Loss/Training Cost = 3411.90

Epoch 19. Loss/Training Cost = 3401.89

Epoch 20. Loss/Training Cost = 3390.25

Epoch 20. Validation Accuracy = 0.46

Epoch 21. Loss/Training Cost = 3380.20

Epoch 22. Loss/Training Cost = 3370.10

Epoch 23. Loss/Training Cost = 3360.46

Epoch 24. Loss/Training Cost = 3352.86

Epoch 25. Loss/Training Cost = 3344.64

Epoch 26. Loss/Training Cost = 3335.97

Epoch 27. Loss/Training Cost = 3328.94

Epoch 28. Loss/Training Cost = 3319.86

Epoch 29. Loss/Training Cost = 3311.74

Epoch 30. Loss/Training Cost = 3305.10

Epoch 30. Validation Accuracy = 0.52

Epoch 31. Loss/Training Cost = 3298.20

Epoch 32. Loss/Training Cost = 3291.37

Epoch 33. Loss/Training Cost = 3285.43

Epoch 34. Loss/Training Cost = 3277.38

Epoch 35. Loss/Training Cost = 3269.67

Epoch 36. Loss/Training Cost = 3264.65

Epoch 37. Loss/Training Cost = 3257.50

Epoch 38. Loss/Training Cost = 3249.14

Epoch 39. Loss/Training Cost = 3244.61

Epoch 40. Loss/Training Cost = 3238.54

Epoch 40. Validation Accuracy = 0.55

Epoch 41. Loss/Training Cost = 3231.89

Epoch 42. Loss/Training Cost = 3226.82

Epoch 43. Loss/Training Cost = 3221.42

Epoch 44. Loss/Training Cost = 3216.49

Epoch 45. Loss/Training Cost = 3211.52

Epoch 46. Loss/Training Cost = 3207.27

Epoch 47. Loss/Training Cost = 3203.28

Epoch 48. Loss/Training Cost = 3197.88

Epoch 49. Loss/Training Cost = 3193.90

Epoch 50. Loss/Training Cost = 3188.47

Epoch 50. Validation Accuracy = 0.57

**SGD Adam data:**

Epoch 0. Loss/Training Cost = 3636.24

Epoch 0. Validation Accuracy = 0.28

Epoch 1. Loss/Training Cost = 3523.90

Epoch 2. Loss/Training Cost = 3364.16

Epoch 3. Loss/Training Cost = 3226.62

Epoch 4. Loss/Training Cost = 3136.25

Epoch 5. Loss/Training Cost = 3074.48

Epoch 6. Loss/Training Cost = 3029.39

Epoch 7. Loss/Training Cost = 2988.72

Epoch 8. Loss/Training Cost = 2954.31

Epoch 9. Loss/Training Cost = 2922.89

Epoch 10. Loss/Training Cost = 2900.52

Epoch 10. Validation Accuracy = 0.70

Epoch 11. Loss/Training Cost = 2879.51

Epoch 12. Loss/Training Cost = 2857.55

Epoch 13. Loss/Training Cost = 2838.86

Epoch 14. Loss/Training Cost = 2823.90

Epoch 15. Loss/Training Cost = 2811.66

Epoch 16. Loss/Training Cost = 2798.65

Epoch 17. Loss/Training Cost = 2787.59

Epoch 18. Loss/Training Cost = 2779.26

Epoch 19. Loss/Training Cost = 2767.85

Epoch 20. Loss/Training Cost = 2758.99

Epoch 20. Validation Accuracy = 0.76

Epoch 21. Loss/Training Cost = 2753.81

Epoch 22. Loss/Training Cost = 2744.53

Epoch 23. Loss/Training Cost = 2740.01

Epoch 24. Loss/Training Cost = 2733.39

Epoch 25. Loss/Training Cost = 2724.32

Epoch 26. Loss/Training Cost = 2717.03

Epoch 27. Loss/Training Cost = 2708.19

Epoch 28. Loss/Training Cost = 2697.89

Epoch 29. Loss/Training Cost = 2691.65

Epoch 30. Loss/Training Cost = 2686.17

Epoch 30. Validation Accuracy = 0.82

Epoch 31. Loss/Training Cost = 2680.15

Epoch 32. Loss/Training Cost = 2674.10

Epoch 33. Loss/Training Cost = 2667.56

Epoch 34. Loss/Training Cost = 2663.95

Epoch 35. Loss/Training Cost = 2658.73

Epoch 36. Loss/Training Cost = 2656.70

Epoch 37. Loss/Training Cost = 2651.12

Epoch 38. Loss/Training Cost = 2644.93

Epoch 39. Loss/Training Cost = 2641.93

Epoch 40. Loss/Training Cost = 2637.69

Epoch 40. Validation Accuracy = 0.85

Epoch 41. Loss/Training Cost = 2635.09

Epoch 42. Loss/Training Cost = 2633.55

Epoch 43. Loss/Training Cost = 2629.44

Epoch 44. Loss/Training Cost = 2627.67

Epoch 45. Loss/Training Cost = 2624.23

Epoch 46. Loss/Training Cost = 2620.23

Epoch 47. Loss/Training Cost = 2617.52

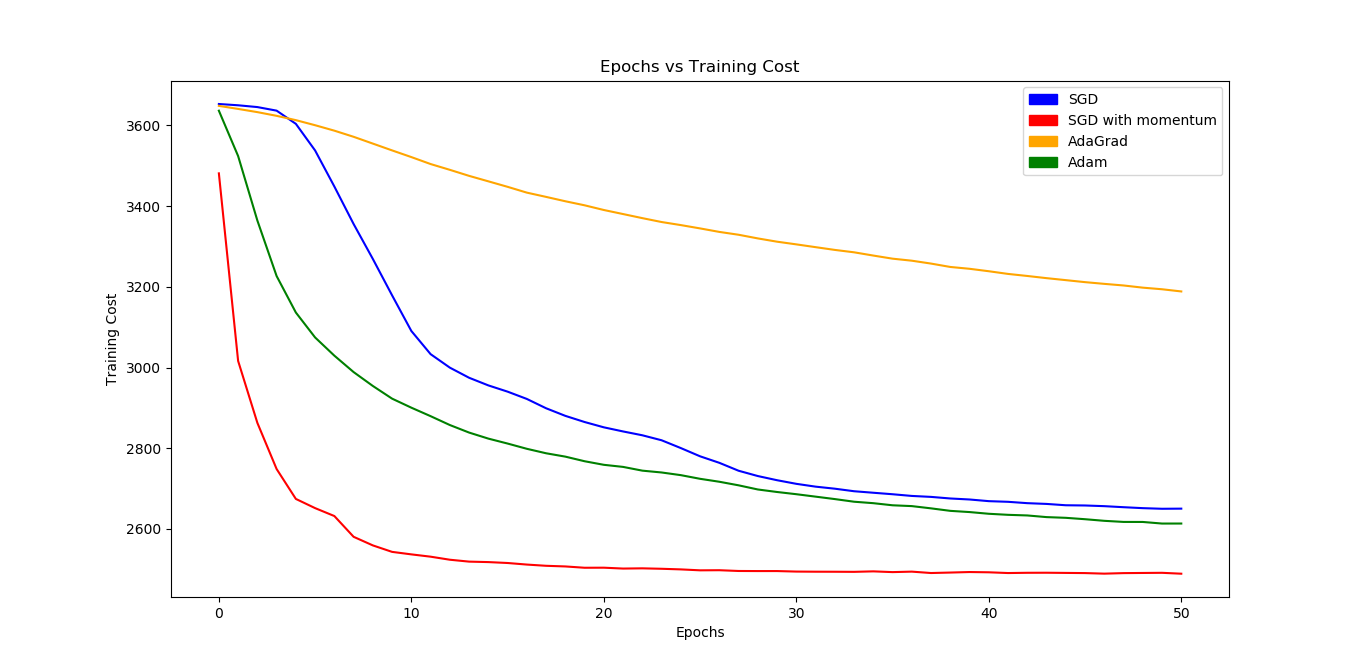
Epoch 48. Loss/Training Cost = 2617.42

Epoch 49. Loss/Training Cost = 2613.48

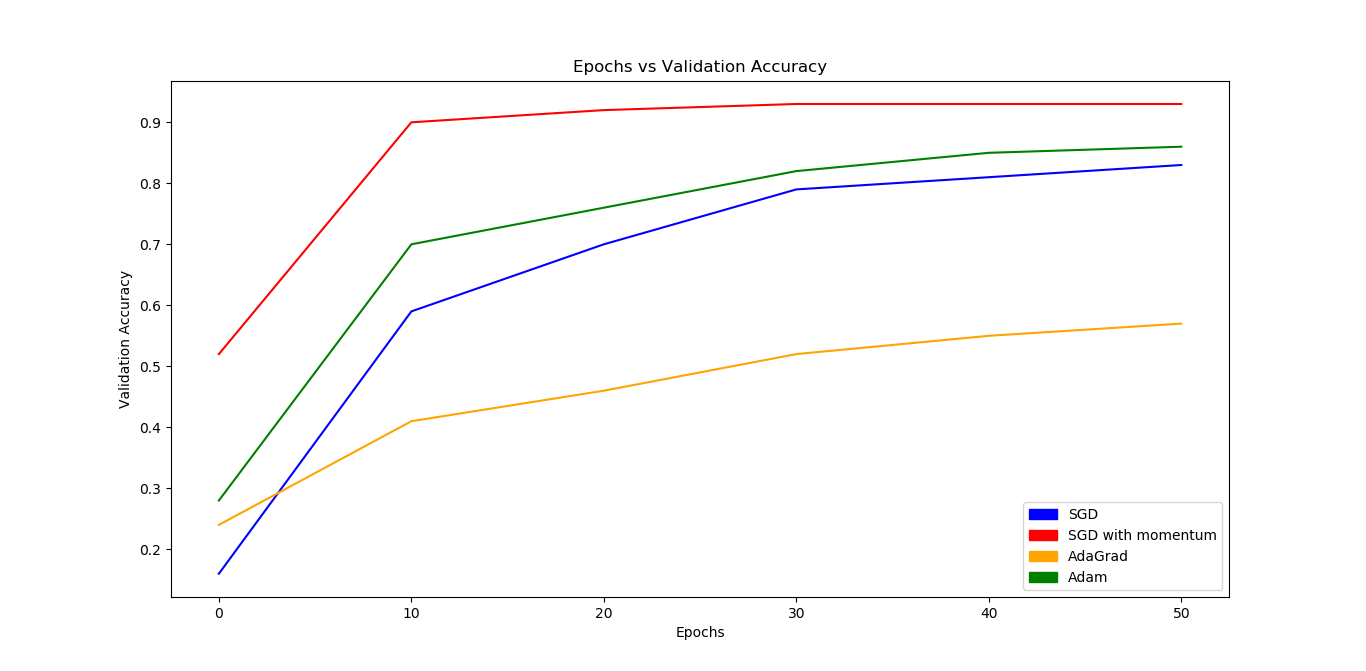
Epoch 50. Loss/Training Cost = 2613.53

Epoch 50. Validation Accuracy = 0.86

1. **Training cost vs number of epochs**

****

1. **Validation accuracy vs number of epochs.**

****

**SGD with momentum = 0.9 seems to be the best approach as it gives the highest accuracy with the least amount of training cost.**

3.2) I applied batch normalization on SGD.

**New learning rate: 0.009**

**Final Test Accuracy: 89%**

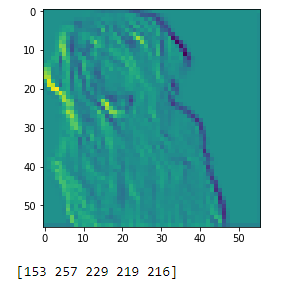
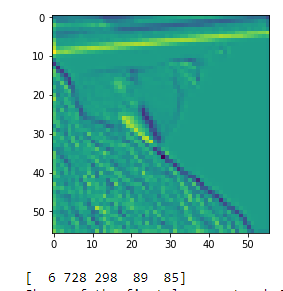
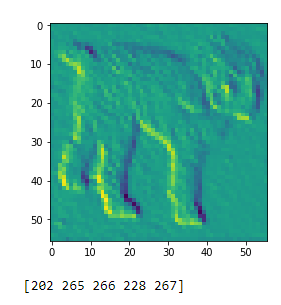
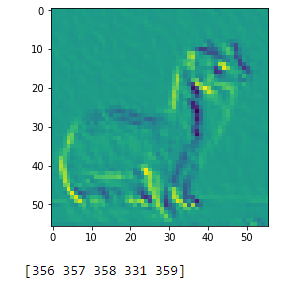
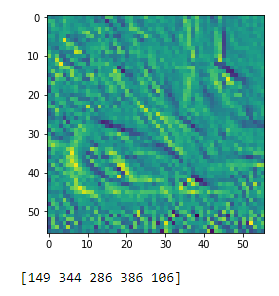
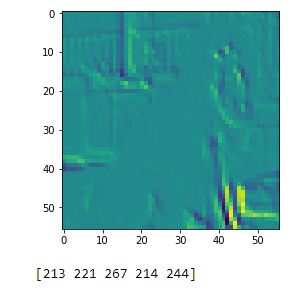
Yes, batch normalization does help.

4.1) *I’ve included the code in prob4.py*

4.2) Shape of the output at the first layer = [64, 56, 56]

‘Plots’ of the output from the first layer:

**Last one is the external image. I’ve included the original image.**



4.3) Shape of the final output layer = [1, 1000]

4.4) Score for external image = [153 257 229 219 216]

**The classification seems correct.**

5.1)

Max-pooling – Since the error is assigned to which it comes from, the other units in the previous layer’s pooling blocks do not contribute to it hence all the others are assigned values of zero.

5.2) *I’ve attached the code in prob5.py*

Epoch 0. Loss/Training Cost = 3582.56

Epoch 0. Validation Accuracy = 0.40

Epoch 1. Loss/Training Cost = 3025.73

Epoch 2. Loss/Training Cost = 2664.89

Epoch 3. Loss/Training Cost = 2596.36

Epoch 4. Loss/Training Cost = 2576.97

Epoch 5. Loss/Training Cost = 2566.53

Epoch 6. Loss/Training Cost = 2558.88

Epoch 7. Loss/Training Cost = 2553.11

Epoch 8. Loss/Training Cost = 2548.00

Epoch 9. Loss/Training Cost = 2543.27

Epoch 10. Loss/Training Cost = 2539.57

Epoch 10. Validation Accuracy = 0.85

Epoch 11. Loss/Training Cost = 2535.89

Epoch 12. Loss/Training Cost = 2532.56

Epoch 13. Loss/Training Cost = 2529.83

Epoch 14. Loss/Training Cost = 2526.71

Epoch 15. Loss/Training Cost = 2524.27

Epoch 16. Loss/Training Cost = 2521.37

Epoch 17. Loss/Training Cost = 2518.88

Epoch 18. Loss/Training Cost = 2516.67

Epoch 19. Loss/Training Cost = 2514.16

Epoch 20. Loss/Training Cost = 2512.22

Epoch 20. Validation Accuracy = 0.87

Epoch 21. Loss/Training Cost = 2510.25

Epoch 22. Loss/Training Cost = 2508.46

Epoch 23. Loss/Training Cost = 2507.02

Epoch 24. Loss/Training Cost = 2505.89

Epoch 25. Loss/Training Cost = 2504.07

Epoch 26. Loss/Training Cost = 2502.93

Epoch 27. Loss/Training Cost = 2501.38

Epoch 28. Loss/Training Cost = 2500.45

Epoch 29. Loss/Training Cost = 2499.41

Epoch 30. Loss/Training Cost = 2498.07

Epoch 30. Validation Accuracy = 0.88

Epoch 31. Loss/Training Cost = 2497.45

Epoch 32. Loss/Training Cost = 2496.29

Epoch 33. Loss/Training Cost = 2495.52

Epoch 34. Loss/Training Cost = 2494.95

Epoch 35. Loss/Training Cost = 2494.01

Epoch 36. Loss/Training Cost = 2493.32

Epoch 37. Loss/Training Cost = 2492.85

Epoch 38. Loss/Training Cost = 2491.85

Epoch 39. Loss/Training Cost = 2491.32

Epoch 40. Loss/Training Cost = 2490.74

Epoch 40. Validation Accuracy = 0.88

Epoch 41. Loss/Training Cost = 2490.14

Epoch 42. Loss/Training Cost = 2489.66

Epoch 43. Loss/Training Cost = 2488.92

Epoch 44. Loss/Training Cost = 2488.61

Epoch 45. Loss/Training Cost = 2488.04

Epoch 46. Loss/Training Cost = 2487.55

Epoch 47. Loss/Training Cost = 2486.93

Epoch 48. Loss/Training Cost = 2486.76

Epoch 49. Loss/Training Cost = 2486.36

Epoch 50. Loss/Training Cost = 2485.85

Epoch 50. Validation Accuracy = 0.89

**Final test accuracy – 89%**

**Design choices:**

The network has two convolutional layers follow by max pooling layers. The final output layer has Softmax as the activation function. I started small and gradually tested on greater number of filters. The 1st convolution layer ended with 20 filters and the 2nd with 64 filters. I used the default PyTorch initialization. It does not use adaptive learning or momentum. I have used L1 regularization and dropout.

**Hyperparameters:**

learning\_rate = 0.005

num\_epochs = 50

batch\_size = 32

dropout\_rate = 0.5

**Based on my results, CNN does perform better than all the other models and optimizers except for SGD with momentum.**