CS541 – Deep Learning
HW – 4

1. Feed – forward Neural Network: -

b) Check grad – The value obtained on running the check grad function is as shown below.

```
2.2005766161823934
```

d) Here the training loss over the final iterations is shown along with the test accuracy obtained with selected best hyper-parameters.

```
Error: 0.5937115469729318  Accuracy: 82.39999999999999 h_star: 5 40 32 0.001 0.0 50 combination: 88
Error: 0.6349604818302009  Accuracy: 82.55 h_star: 5 40 32 0.005 0.0 50 combination: 89
Error: nan  Accuracy: 10.7 h_star: 5 40 32 0.01 0.0 50 combination: 90
Error: nan  Accuracy: 10.7 h_star: 5 40 32 0.1 0.0 50 combination: 91
Error: 0.6155291210972825  Accuracy: 82.1 h_star: 5 40 64 0.001 0.0 50 combination: 92
Error: 0.6217796525270429  Accuracy: 79.25 h_star: 5 40 64 0.005 0.0 50 combination: 93
Error: 2.3043456574843586  Accuracy: 8.649999999999999 h_star: 5 40 64 0.01 0.0 50 combination: 94
Error: nan  Accuracy: 10.7 h_star: 5 40 64 0.1 0.0 50 combination: 95
Error: 0.5884153042336675  Accuracy: 81.5 h_star: 5 50 16 0.001 0.0 50 combination: 96
Error: 0.6808991860132744  Accuracy: 82.45 h_star: 5 50 16 0.005 0.0 50 combination: 97
Error: 0.6439063227065468  Accuracy: 81.95 h_star: 5 50 16 0.01 0.0 50 combination: 98
Error: 2.30639836467126  Accuracy: 8.649999999999999 h_star: 5 50 16 0.1 0.0 50 combination: 99
Error: 0.6210838592055883  Accuracy: 82.15 h_star: 5 50 32 0.001 0.0 50 combination: 100
Error: 0.6409867703058483  Accuracy: 82.85 h_star: 5 50 32 0.005 0.0 50 combination: 101
Error: nan  Accuracy: 10.7 h_star: 5 50 32 0.01 0.0 50 combination: 102
Error: nan  Accuracy: 10.7 h_star: 5 50 32 0.1 0.0 50 combination: 103
Error: 0.5778338179400714  Accuracy: 81.75 h_star: 5 50 64 0.001 0.0 50 combination: 104
Error: 0.5413918535366734  Accuracy: 82.39999999999999 h_star: 5 50 64 0.005 0.0 50 combination: 105
Error: 2.3048281843017695  Accuracy: 9.700000000000001 h_star: 5 50 64 0.01 0.0 50 combination: 106
Error: nan  Accuracy: 10.7 h_star: 5 50 64 0.1 0.0 50 combination: 107
```

```
Best Hyperparameters:
32 0.001 0.0 50
FINAL Validation (0.6210838592055883, 82.15)



FINAL TESTING ERROR:  0.5712772994721764
Accuracy:  82.83166666666666
```

We achieved an accuracy of 82.831% on the test dataset for the following best hyper-parameters.

NUM_HIDDEN_LAYERS = 3
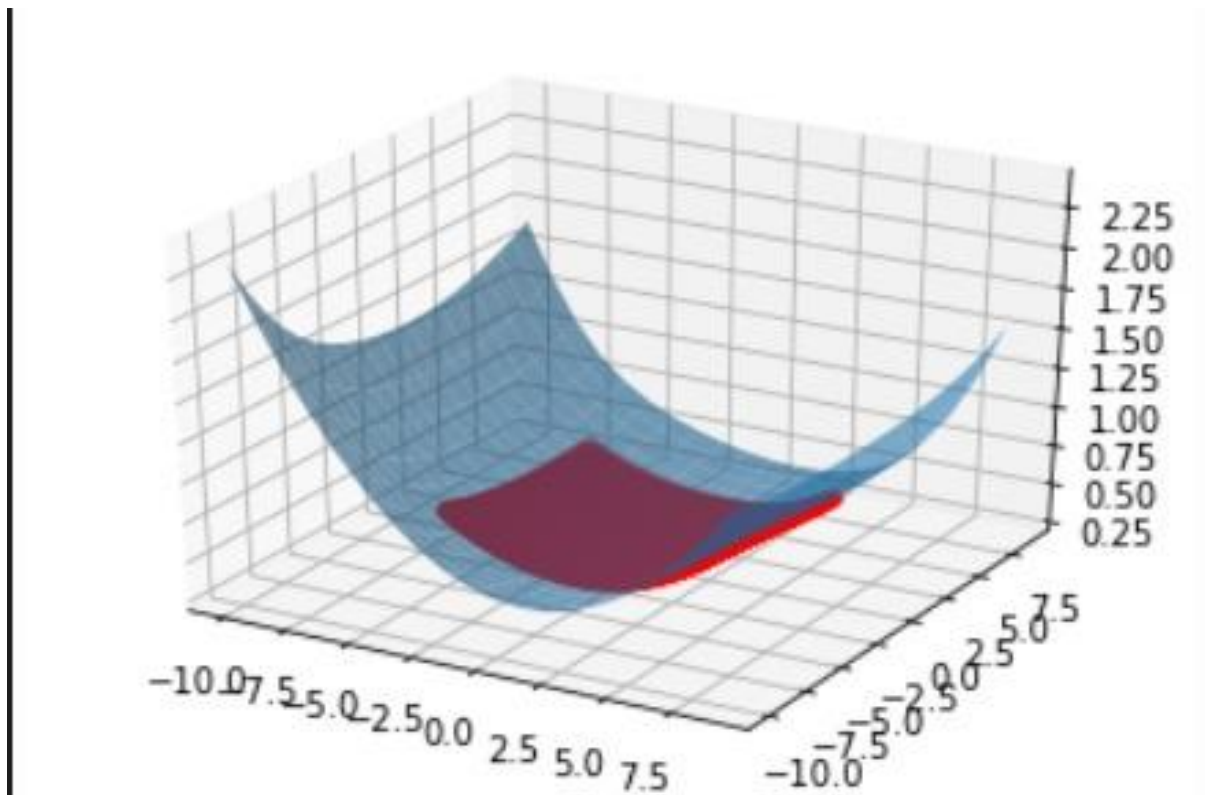NUM_HIDDEN = 50
Batch Size = 32
Epochs = 50
Learning Rate = 0.001
Regularization Strength = 0
Alpha = 0.0

2. Mountains and Valleys: -
a) and b) – Plot of training fce (Cross-Entropy Loss function) against the two most varying parameters. The fce is plotted along the z-axis for all the values of the training data.

The training loss (fce) can be seen to reduce from around 2.12 to 0 .25

Done by –
1) Shrishailya Chavan
2) Uday Sankar
3) Gowri Shankar Sai Manikandan