

ASSIGNMENT 3

CS5691 Pattern Recognition and Machine Learning

CS5691 Assignment 3

Team Members:

BE17B007 N Sowmya Manojna
PH17B010 Thakkar Riya Anandbhai
PH17B011 Chaithanya Krishna Moorthy

Indian Institute of Technology, Madras



Contents

1	Dataset 1A	2
1.1	Perceptron	2
1.2	MLFFNN	2
1.2.1	Classification Accuracies	2
1.2.2	Best Model	2
1.2.3	Decision Region	3
1.3	Linear SVM	4
2	Dataset 1B	5
2.1	MLFFNN	5
2.1.1	Classification Accuracies	5
2.1.2	Best Model	5
2.1.3	Decision Region	6
2.1.4	Surface Plots	6
2.1.4.1	Hidden Layer 1, Node 1	7
2.1.4.2	Hidden Layer 1, Node 2	8
2.1.4.3	Hidden Layer 1, Node 3	9
2.1.4.4	Hidden Layer 1, Node 4	10
2.1.4.5	Hidden Layer 1, Node 5	11
2.1.4.6	Hidden Layer 1, Node 6	12
2.1.4.7	Hidden Layer 1, Node 7	13
2.1.4.8	Hidden Layer 1, Node 8	14
2.1.4.9	Hidden Layer 2, Node 1	15
2.1.4.10	Hidden Layer 2, Node 2	16
2.1.4.11	Hidden Layer 2, Node 3	17
2.1.4.12	Hidden Layer 2, Node 4	18
2.1.4.13	Hidden Layer 2, Node 5	19
2.1.4.14	Hidden Layer 2, Node 6	20
2.1.4.15	Hidden Layer 2, Node 7	21
2.1.4.16	Hidden Layer 2, Node 8	22
2.1.4.17	Output Layer, Node 1	23
2.1.4.18	Output Layer, Node 2	24
2.1.4.19	Output Layer, Node 3	25
2.2	Non-Linear SVM	26
3	Dataset 2A	27
3.1	MLFFNN	27
3.2	Gaussian-kernel SVM	27

1 Dataset 1A

1.1 Perceptron

1.2 MLFFNN

1.2.1 Classification Accuracies

The classification accuracies on the training and validation datasets are as follows:

# Neurons	Activation	Solver	Batch Size	α	Learning Rate	Accuracy	Validation Accuracy
5	tanh	lbfgs	200	0.0001	adaptive	100.0	100.0
5	tanh	lbfgs	200	0.0001	constant	100.0	100.0
5	tanh	lbfgs	200	0.0	invscaling	100.0	100.0
5	tanh	lbfgs	200	0.0	adaptive	100.0	100.0
5	tanh	lbfgs	200	0.0	constant	100.0	100.0
5	tanh	lbfgs	100	0.0	adaptive	100.0	100.0
5	tanh	lbfgs	100	0.0001	invscaling	100.0	100.0
5	relu	lbfgs	200	0.0	constant	100.0	100.0
5	relu	lbfgs	100	0.0001	invscaling	100.0	100.0
5	relu	lbfgs	200	0.0	adaptive	100.0	100.0

Table 2: Best 10 Train and Validation Accuracies obtained after performing a [GridSearch](#) on 432 parameter combinations.

1.2.2 Best Model

The parameter combination were additionally sorted based on minimum fitting time (least fitting time - first) and the model that gave the best accuracy the fastest was chosen. Hence the best parameter combination chosen is:

- hidden_layer_sizes: 5
- activation: tanh
- solver: lbfgs
- batch_size: 200
- alpha: 0.0001
- learning_rate: adaptive

The classification accuracy of the best model on the testing data is: 100%. The confusion matrices obtained are as follows:

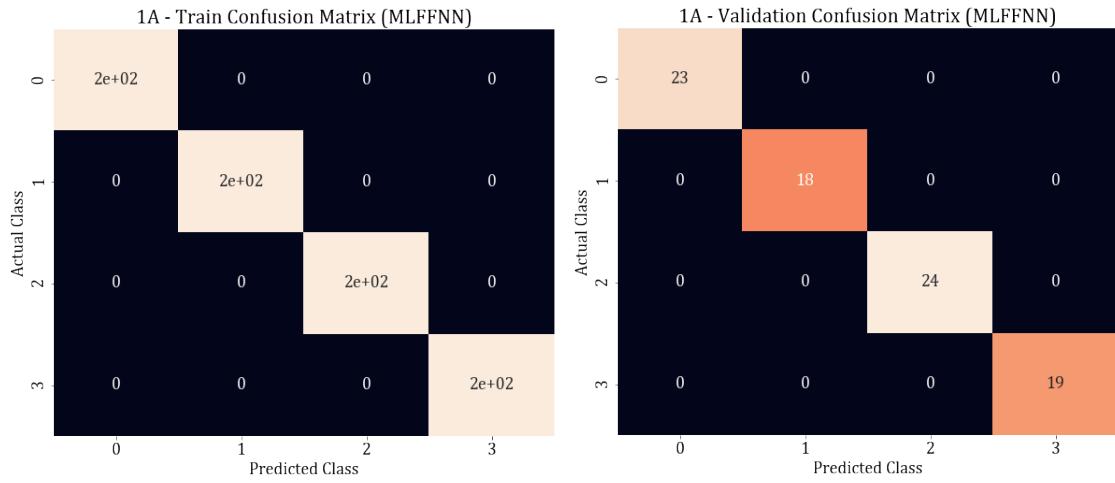


Figure 1: Training and Validation confusion matrices obtained for the best parameter combination, on the left and right respectively.

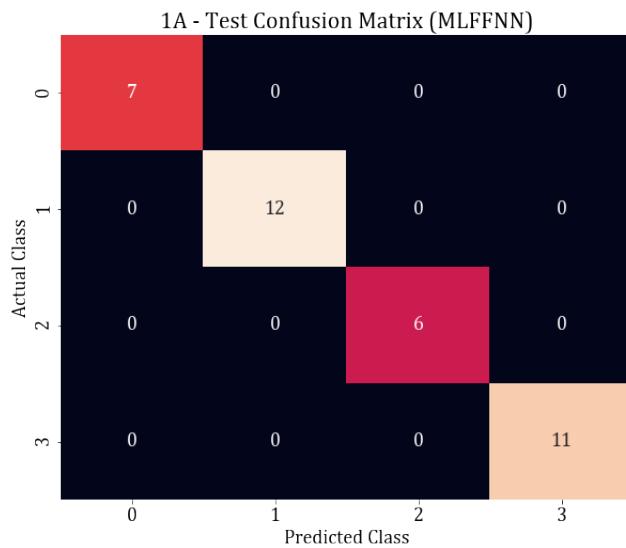


Figure 2: Testing confusion matrices obtained for the best parameter combination.

1.2.3 Decision Region

The decision region plots obtained is as follows:

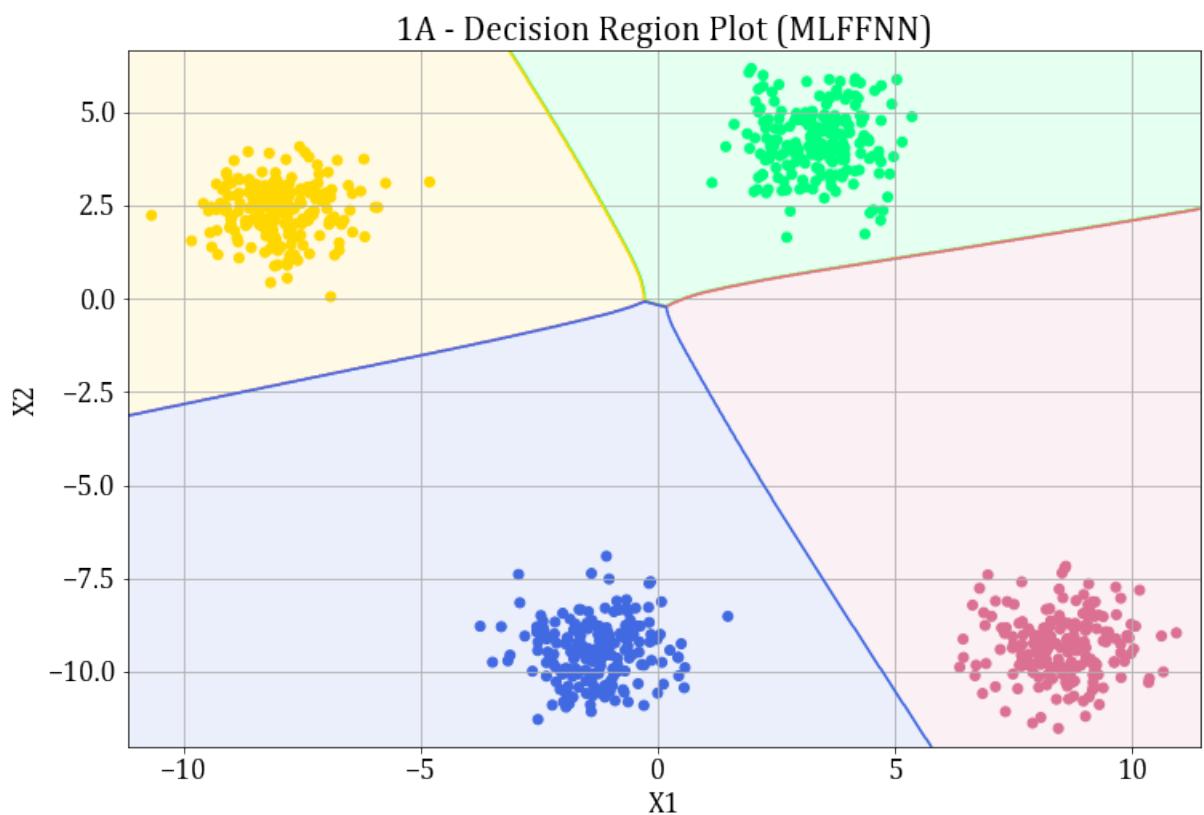


Figure 3: Decision Region Plot obtained for the best parameter combination.

1.3 Linear SVM

2 Dataset 1B

2.1 MLFFNN

2.1.1 Classification Accuracies

The classification accuracies on the training and validation datasets are as follows:

# Neurons	Activation	Batch Size	Early Stopping	Learning Rate	α	Accuracy	Validation Accuracy
(8, 8)	relu	50	False	adaptive	0.01	99.33	98.41
(8, 8)	relu	50	False	constant	0.001	99.33	98.41
(8, 8)	relu	50	False	invscaling	0.01	99.33	98.41
(8, 8)	relu	50	False	adaptive	0.001	99.33	98.41
(8, 8)	relu	50	False	invscaling	0.001	99.33	98.41
(8, 8)	relu	50	False	constant	0.01	99.33	98.41
(10, 10)	relu	50	False	adaptive	0.01	99.0	98.41
(10, 10)	relu	50	False	constant	0.01	99.0	98.41
(10, 10)	relu	50	False	invscaling	0.01	99.0	98.41
(10, 10)	relu	50	False	constant	0.001	99.0	96.82

Table 3: Best 10 Train and Validation Accuracies obtained after performing a [GridSearch](#) on 432 parameter combinations.

2.1.2 Best Model

The parameter combination were additionally sorted based on minimum fitting time (least fitting time - first) and the model that gave the best accuracy the fastest was chosen. Hence the best parameter combination chosen is:

- hidden_layer_sizes: (8, 8)
- activation: relu
- batch_size: 50
- early_stopping: False
- learning_rate: adaptive
- alpha: 0.01

The classification accuracy of the best model on the testing data is: 96.296%. The confusion matrices obtained are as follows:

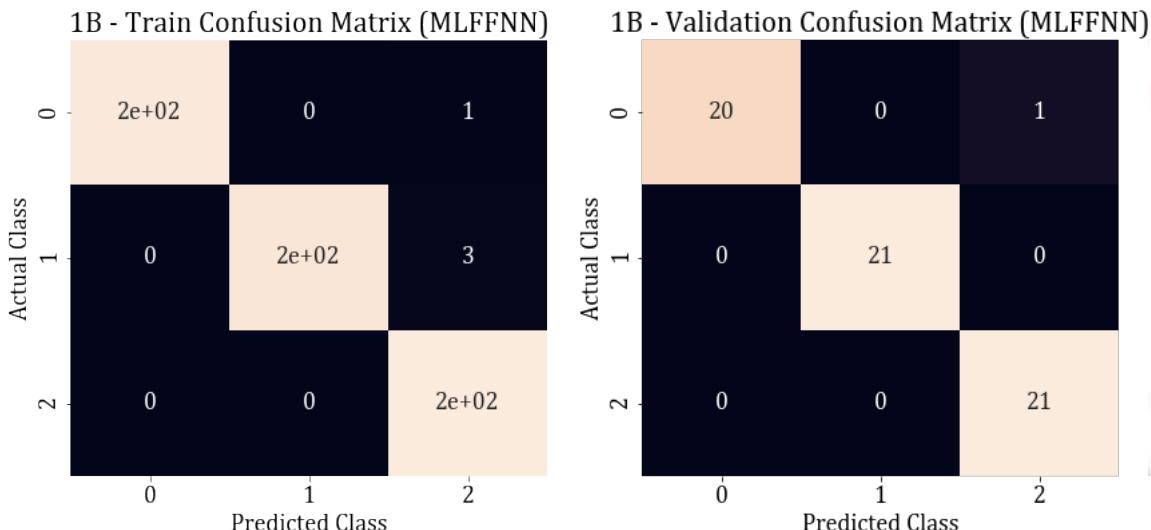


Figure 4: Training and Validation confusion matrices obtained for the best parameter combination, on the left and right respectively.

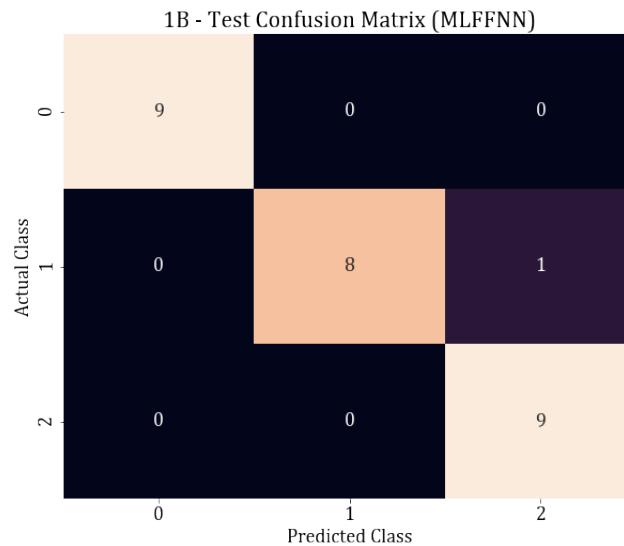


Figure 5: Testing confusion matrices obtained for the best parameter combination.

2.1.3 Decision Region

The decision region plots obtained is as follows:

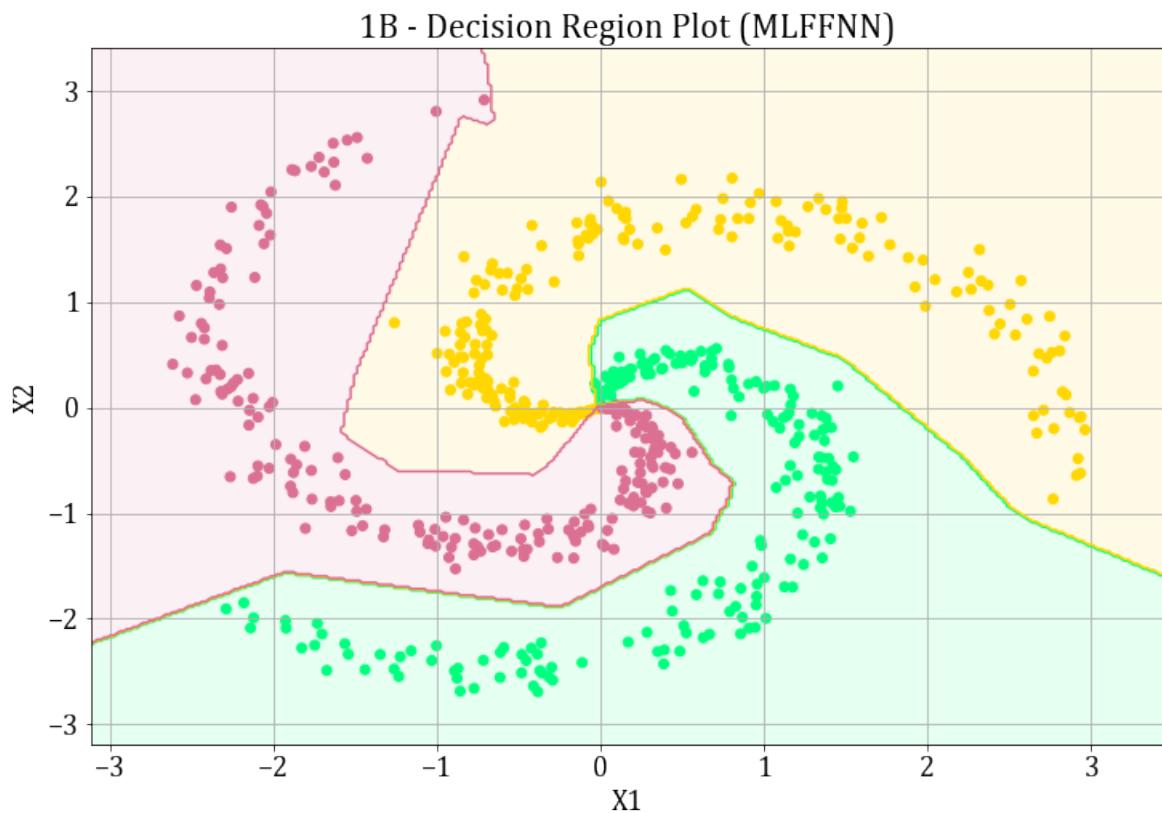


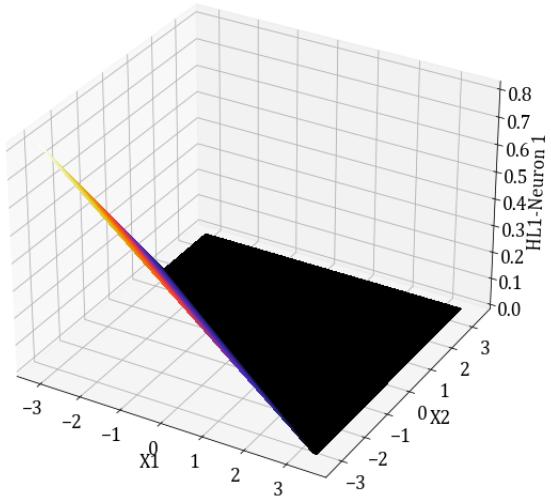
Figure 6: Decision Region Plot obtained for the best parameter combination.

2.1.4 Surface Plots

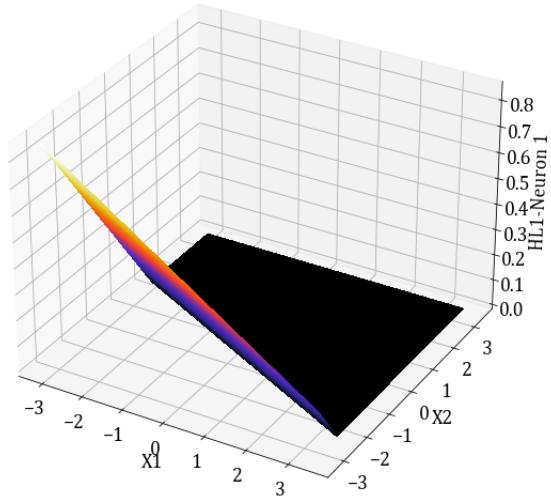
The neuron-wise surface plots obtained for the hidden and output layers is as follows:

2.1.4.1 Hidden Layer 1, Node 1

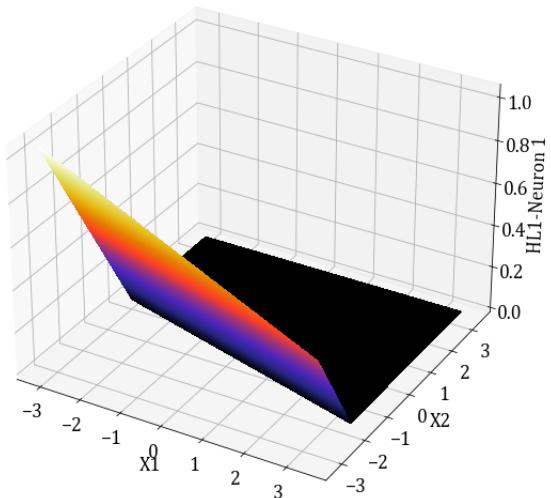
Epoch: 1; Surface for Layer 1, Neuron 1



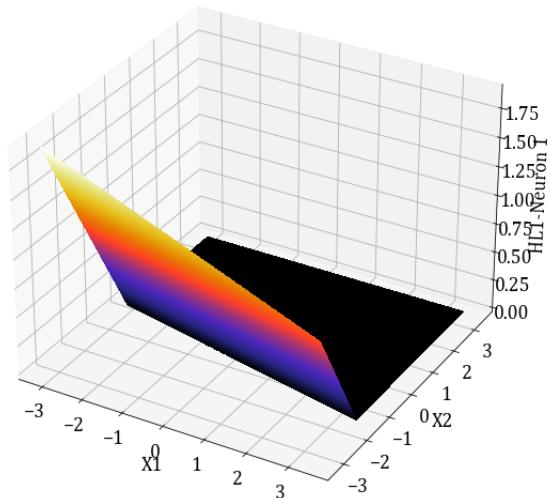
Epoch: 5; Surface for Layer 1, Neuron 1



Epoch: 20; Surface for Layer 1, Neuron 1



Epoch: 100; Surface for Layer 1, Neuron 1



Converged; Surface for Layer 1, Neuron 1

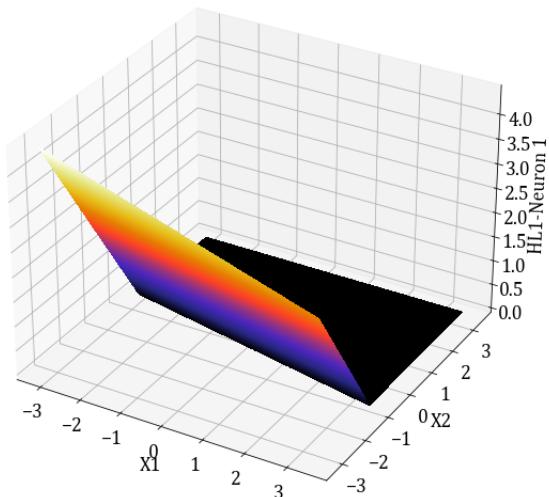
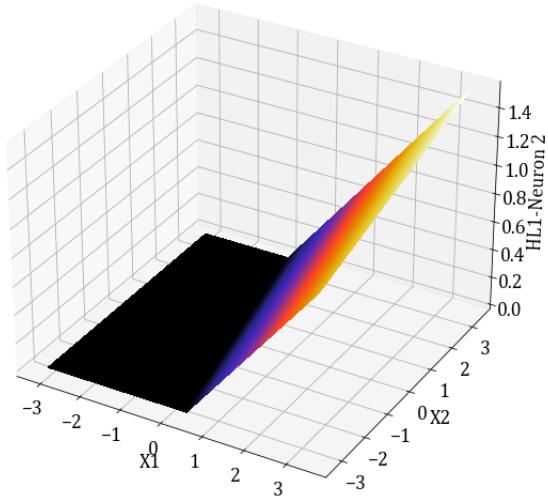


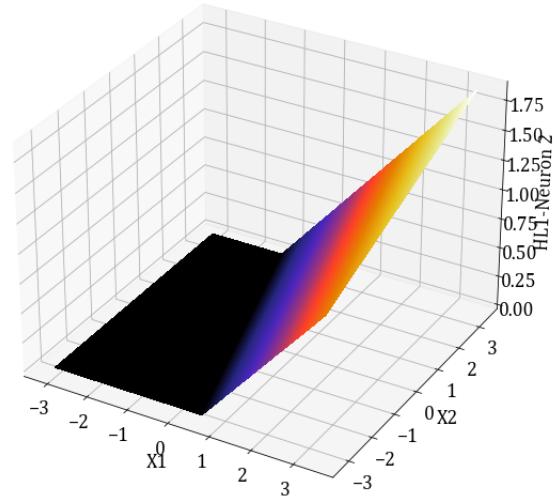
Figure 7: Surface Plots obtained for Hidden Layer 1, Neuron 1, across epochs.

2.1.4.2 Hidden Layer 1, Node 2

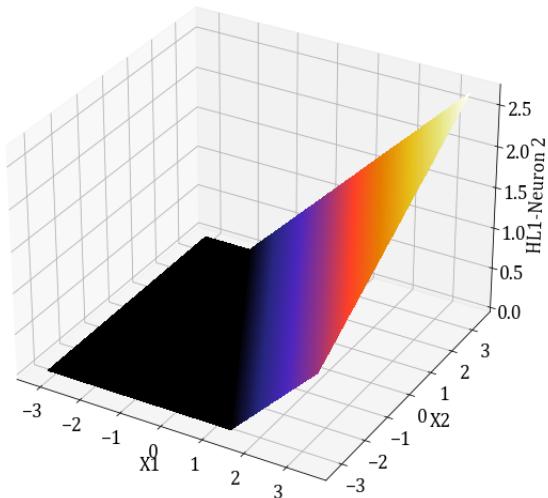
Epoch: 1; Surface for Layer 1, Neuron 2



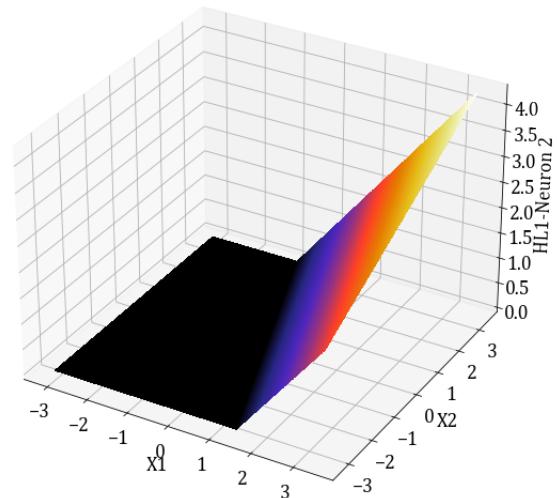
Epoch: 5; Surface for Layer 1, Neuron 2



Epoch: 20; Surface for Layer 1, Neuron 2



Epoch: 100; Surface for Layer 1, Neuron 2



Converged; Surface for Layer 1, Neuron 2

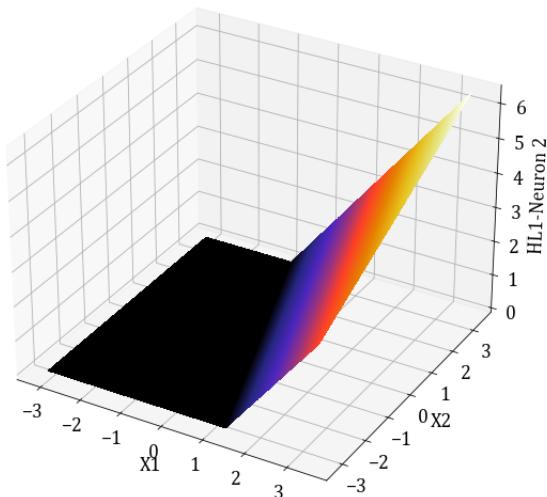
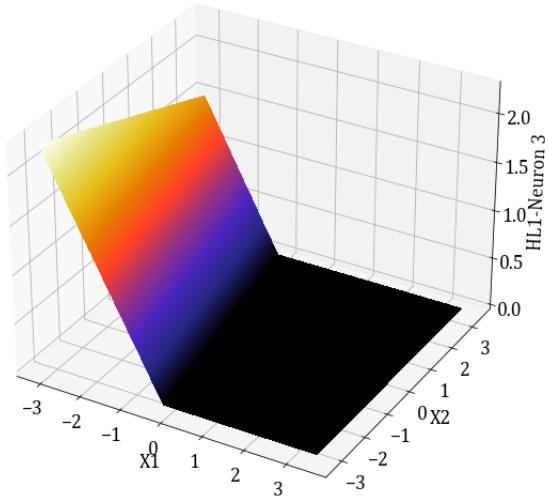


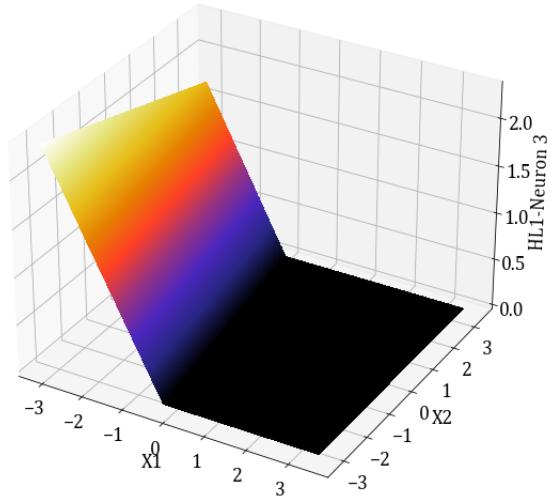
Figure 8: Surface Plots obtained for Hidden Layer 1, Neuron 2, across epochs.

2.1.4.3 Hidden Layer 1, Node 3

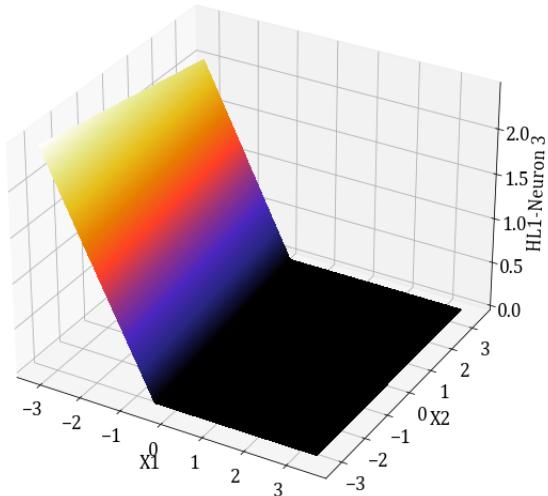
Epoch: 1; Surface for Layer 1, Neuron 3



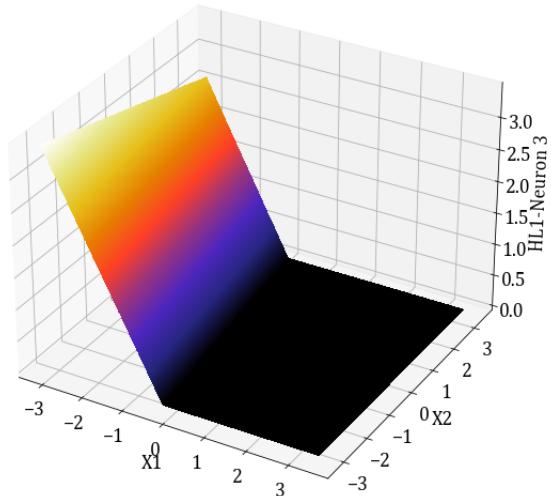
Epoch: 5; Surface for Layer 1, Neuron 3



Epoch: 20; Surface for Layer 1, Neuron 3



Epoch: 100; Surface for Layer 1, Neuron 3



Converged; Surface for Layer 1, Neuron 3

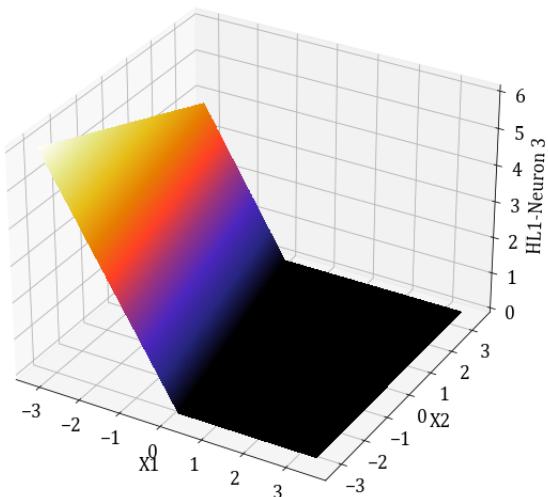


Figure 9: Surface Plots obtained for Hidden Layer 1, Neuron 3, across epochs.

2.1.4.4 Hidden Layer 1, Node 4

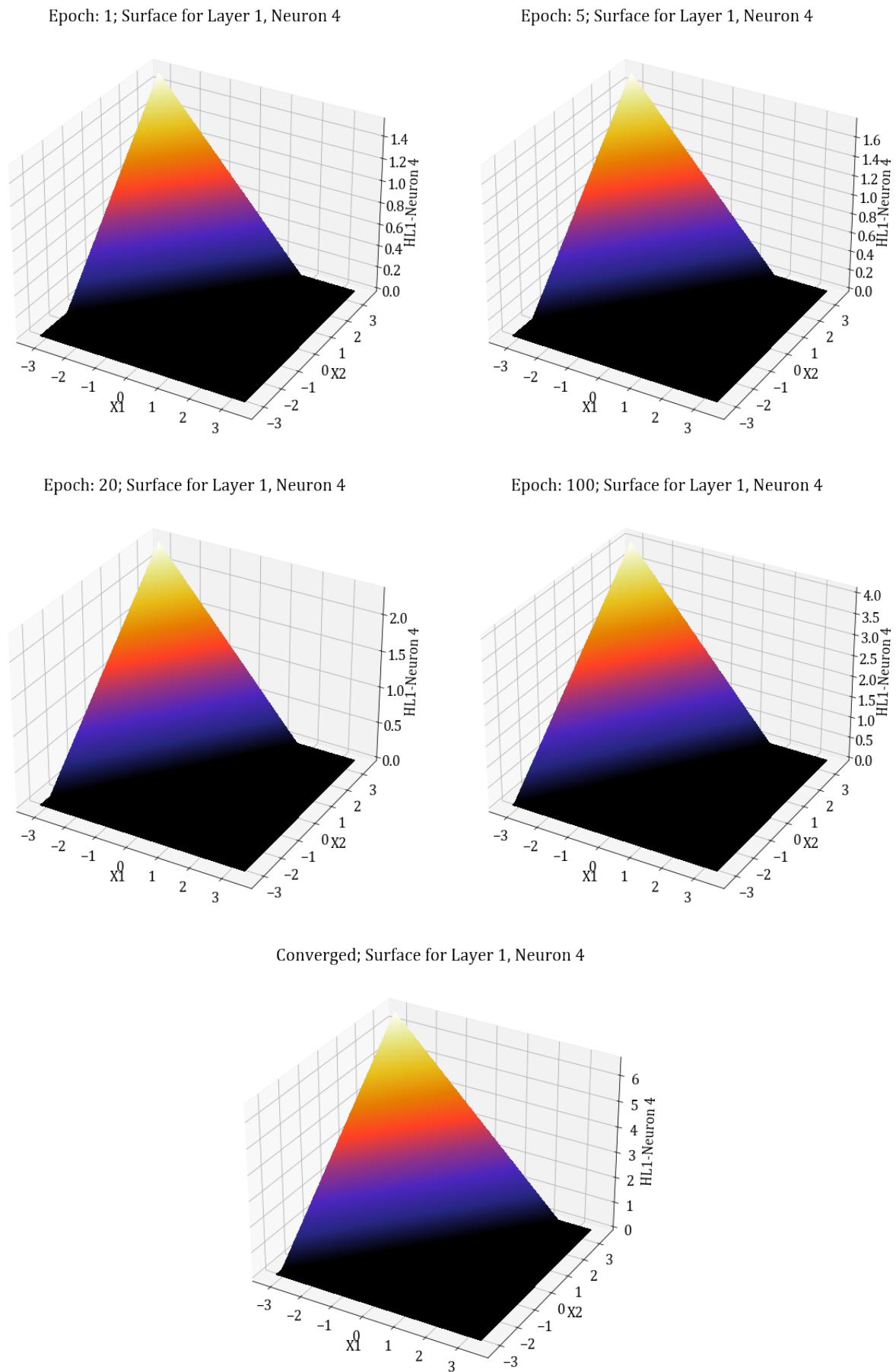
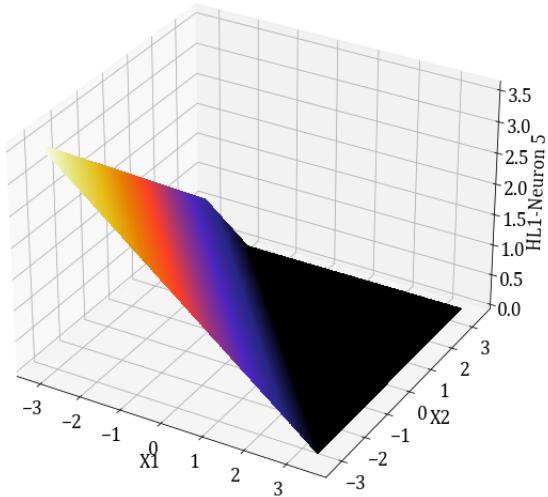


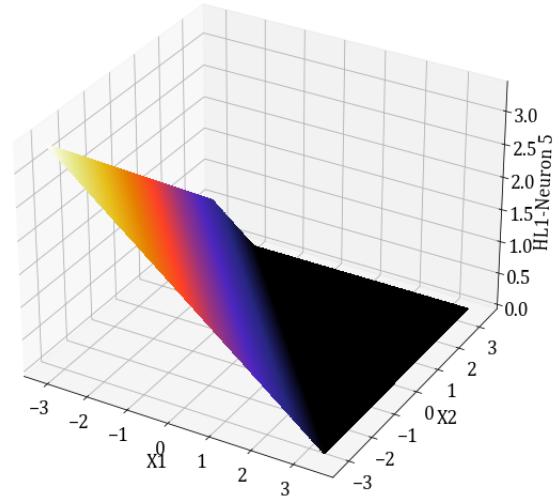
Figure 10: Surface Plots obtained for Hidden Layer 1, Neuron 4, across epochs.

2.1.4.5 Hidden Layer 1, Node 5

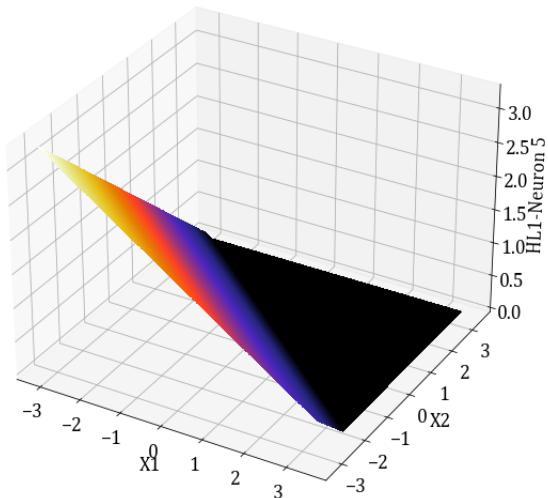
Epoch: 1; Surface for Layer 1, Neuron 5



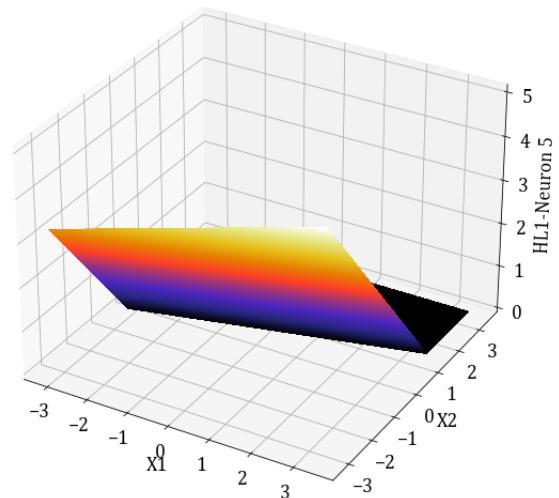
Epoch: 5; Surface for Layer 1, Neuron 5



Epoch: 20; Surface for Layer 1, Neuron 5



Epoch: 100; Surface for Layer 1, Neuron 5



Converged; Surface for Layer 1, Neuron 5

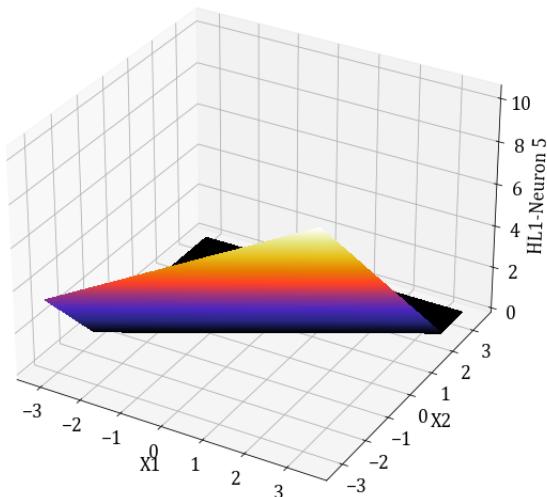


Figure 11: Surface Plots obtained for Hidden Layer 1, Neuron 5, across epochs.

2.1.4.6 Hidden Layer 1, Node 6

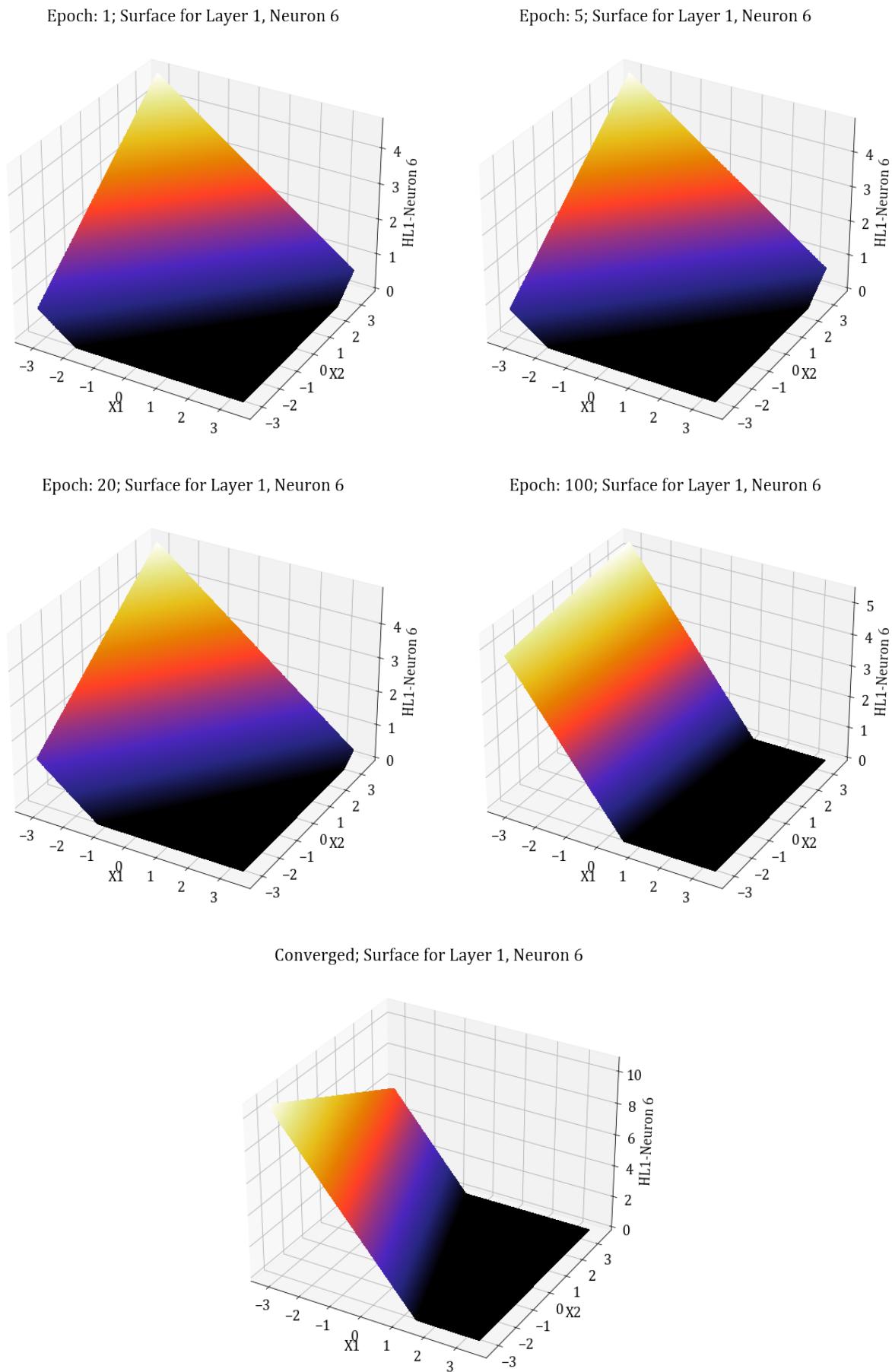
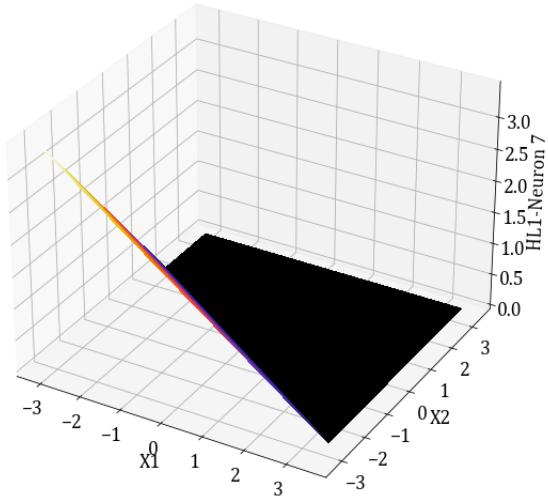


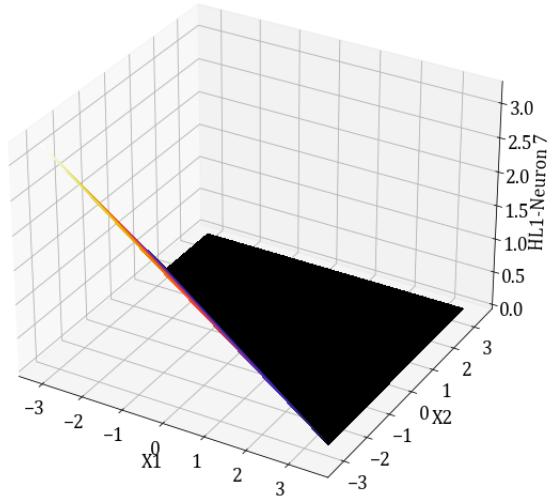
Figure 12: Surface Plots obtained for Hidden Layer 1, Neuron 6, across epochs.

2.1.4.7 Hidden Layer 1, Node 7

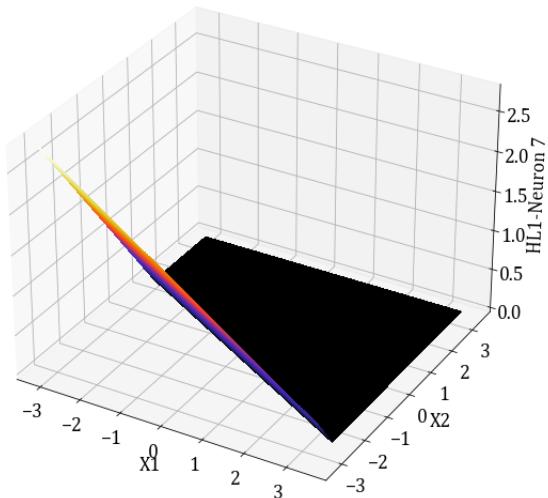
Epoch: 1; Surface for Layer 1, Neuron 7



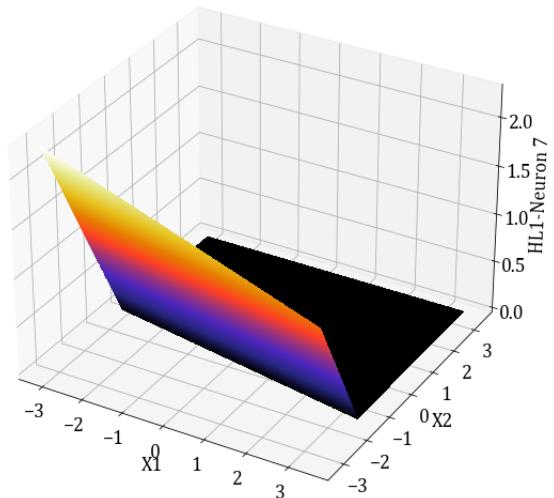
Epoch: 5; Surface for Layer 1, Neuron 7



Epoch: 20; Surface for Layer 1, Neuron 7



Epoch: 100; Surface for Layer 1, Neuron 7



Converged; Surface for Layer 1, Neuron 7

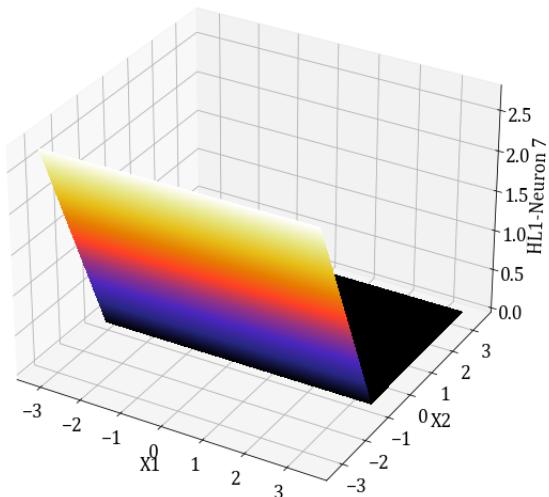


Figure 13: Surface Plots obtained for Hidden Layer 1, Neuron 7, across epochs.

2.1.4.8 Hidden Layer 1, Node 8

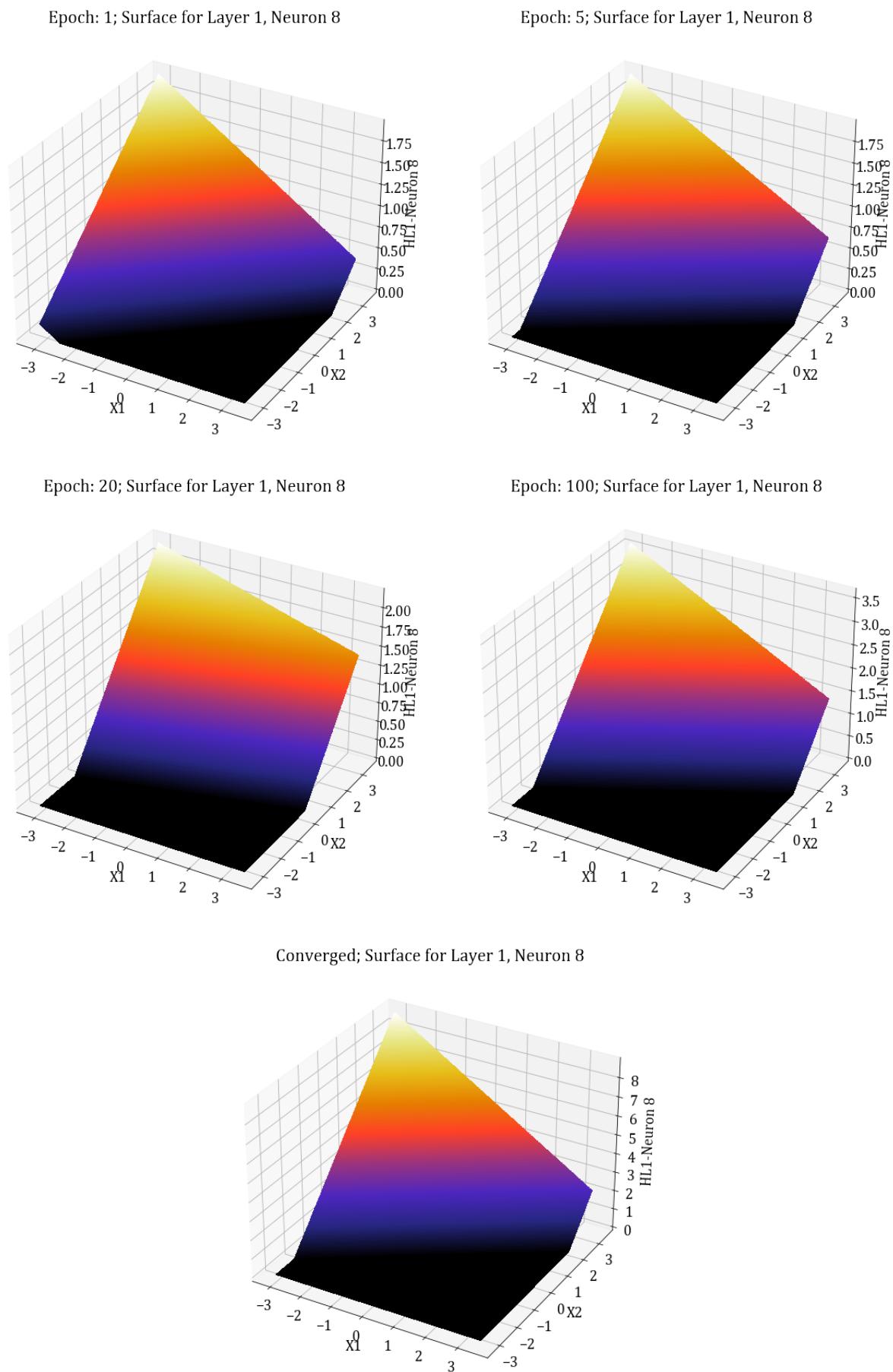
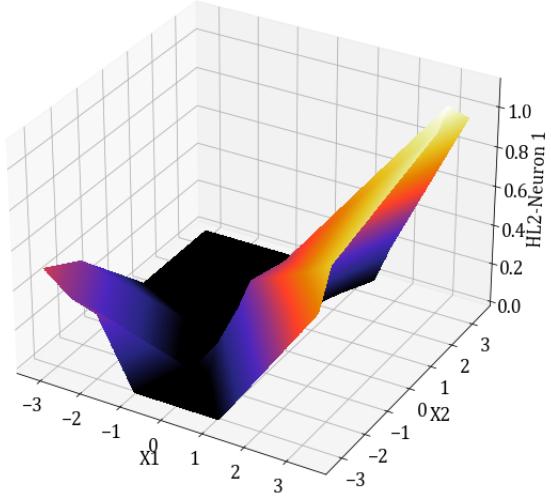


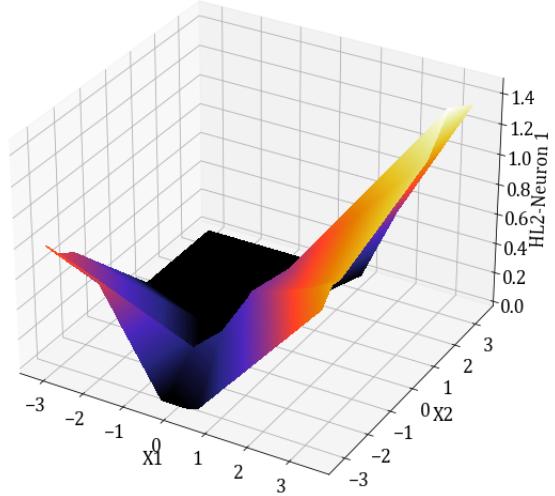
Figure 14: Surface Plots obtained for Hidden Layer 1, Neuron 8, across epochs.

2.1.4.9 Hidden Layer 2, Node 1

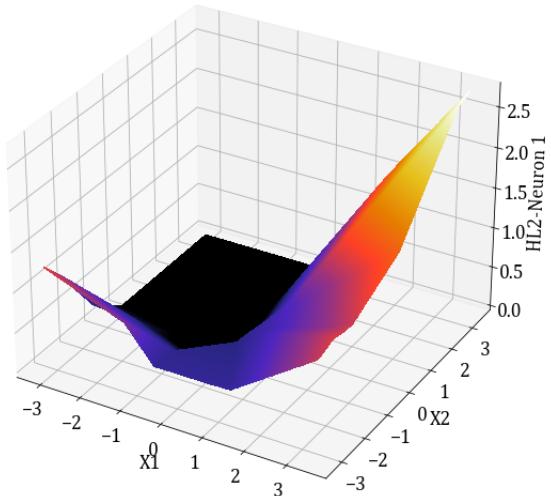
Epoch: 1; Surface for Layer 2, Neuron 1



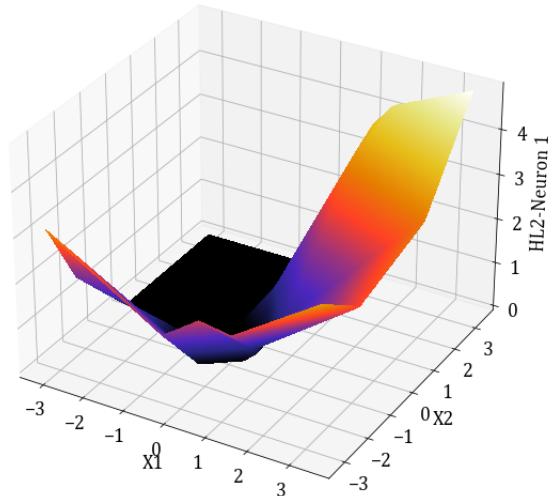
Epoch: 5; Surface for Layer 2, Neuron 1



Epoch: 20; Surface for Layer 2, Neuron 1



Epoch: 100; Surface for Layer 2, Neuron 1



Converged; Surface for Layer 2, Neuron 1

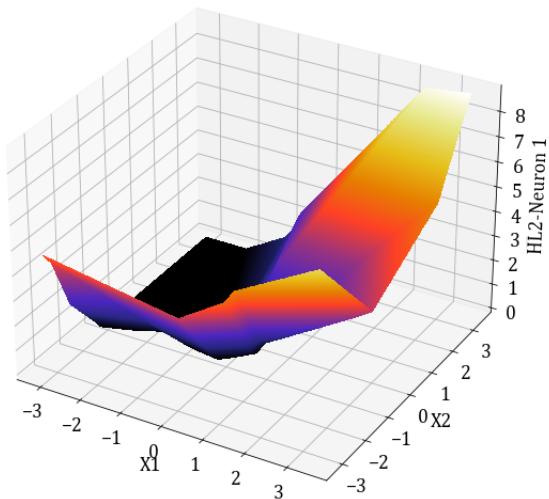
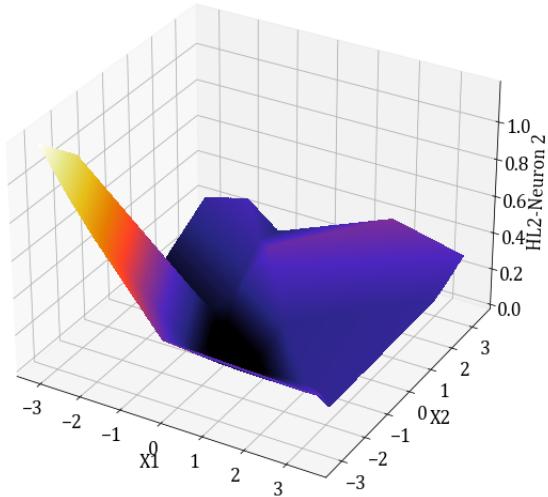


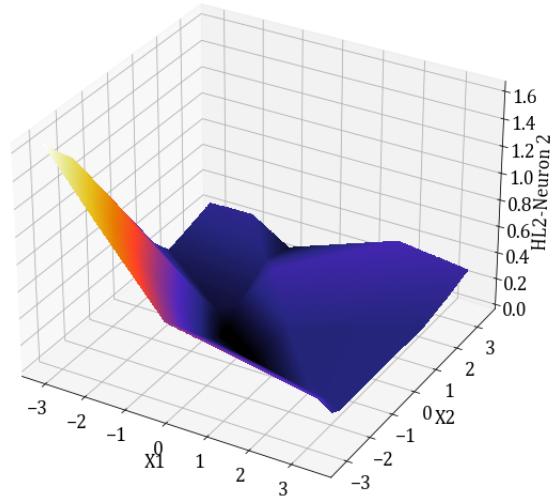
Figure 15: Surface Plots obtained for Hidden Layer 2, Neuron 1, across epochs.

2.1.4.10 Hidden Layer 2, Node 2

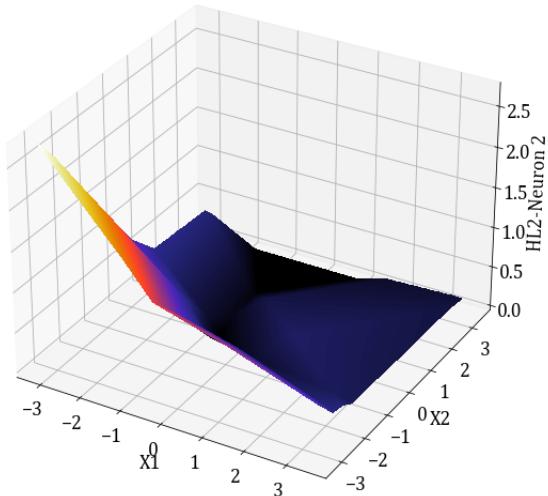
Epoch: 1; Surface for Layer 2, Neuron 2



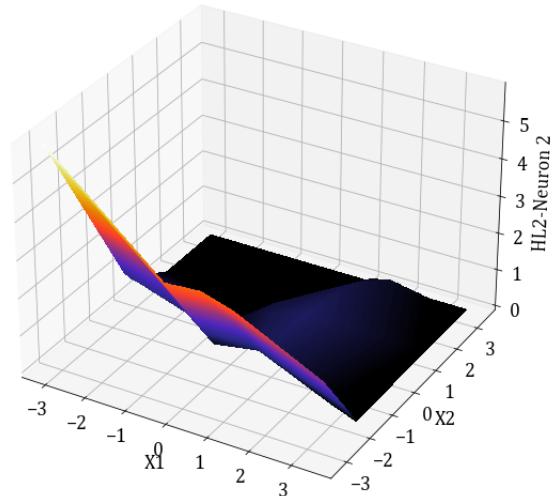
Epoch: 5; Surface for Layer 2, Neuron 2



Epoch: 20; Surface for Layer 2, Neuron 2



Epoch: 100; Surface for Layer 2, Neuron 2



Converged; Surface for Layer 2, Neuron 2

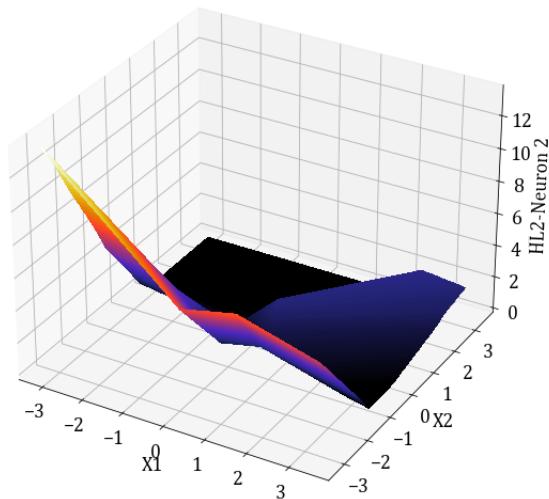
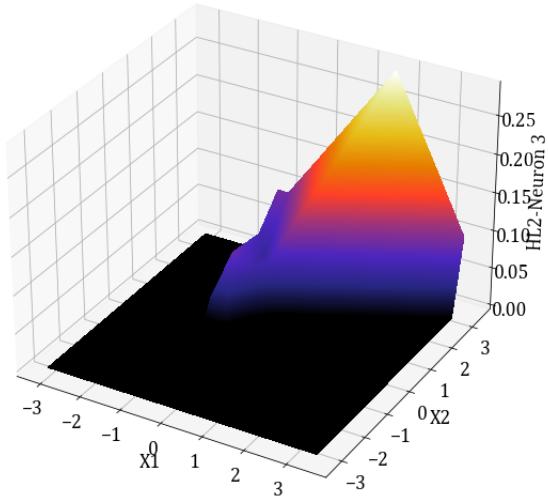


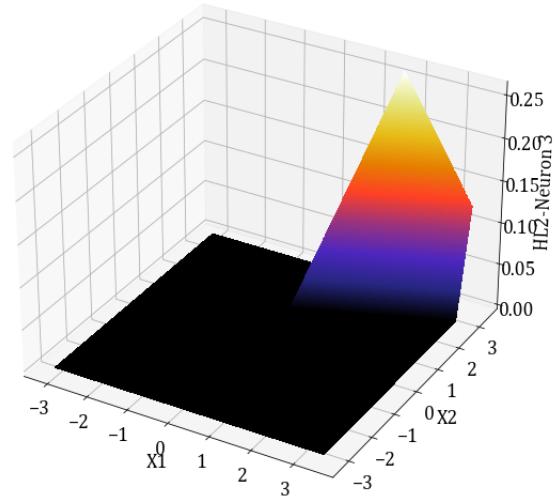
Figure 16: Surface Plots obtained for Hidden Layer 2, Neuron 2, across epochs.

2.1.4.11 Hidden Layer 2, Node 3

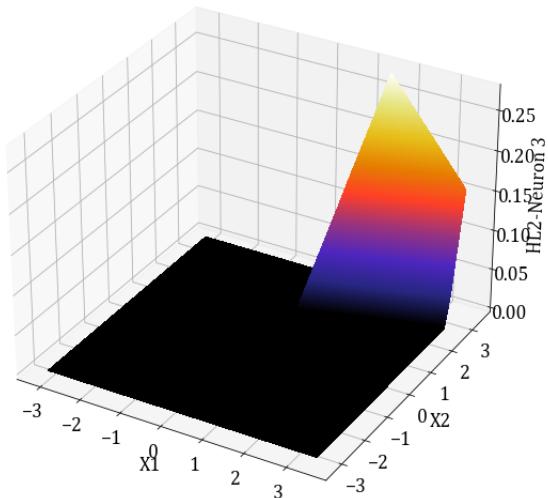
Epoch: 1; Surface for Layer 2, Neuron 3



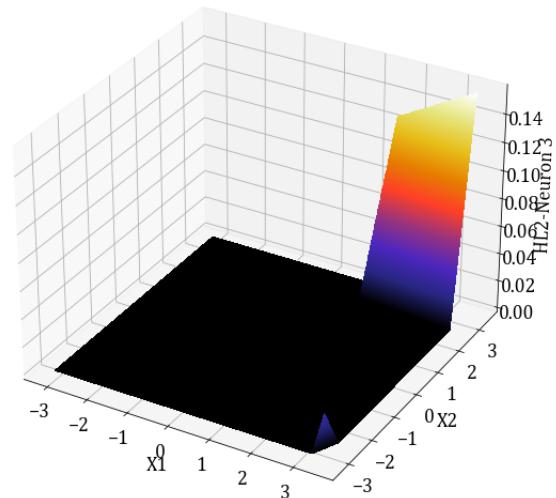
Epoch: 5; Surface for Layer 2, Neuron 3



Epoch: 20; Surface for Layer 2, Neuron 3



Epoch: 100; Surface for Layer 2, Neuron 3



Converged; Surface for Layer 2, Neuron 3

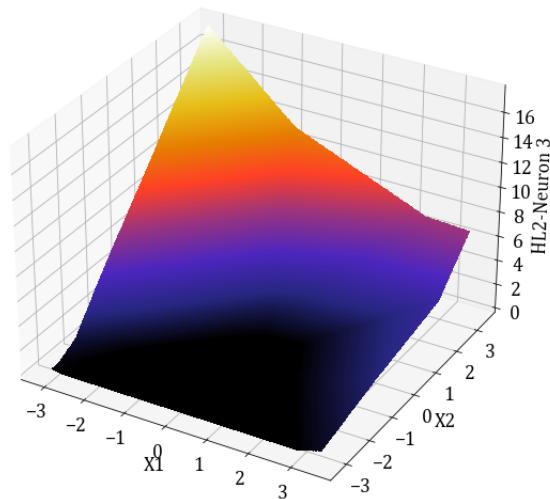
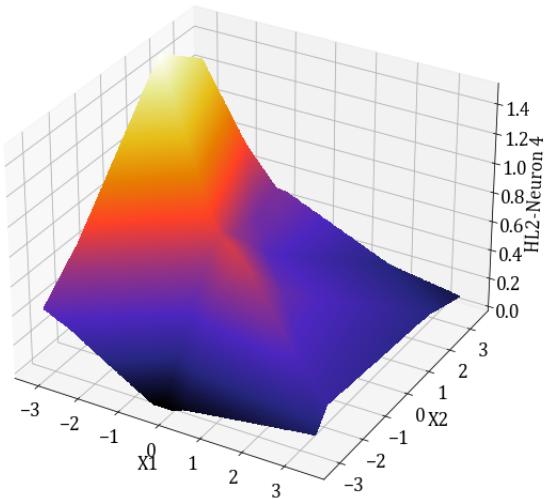


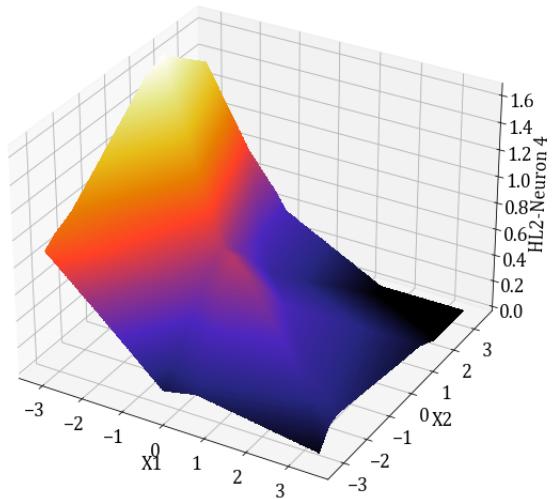
Figure 17: Surface Plots obtained for Hidden Layer 2, Neuron 3, across epochs.

2.1.4.12 Hidden Layer 2, Node 4

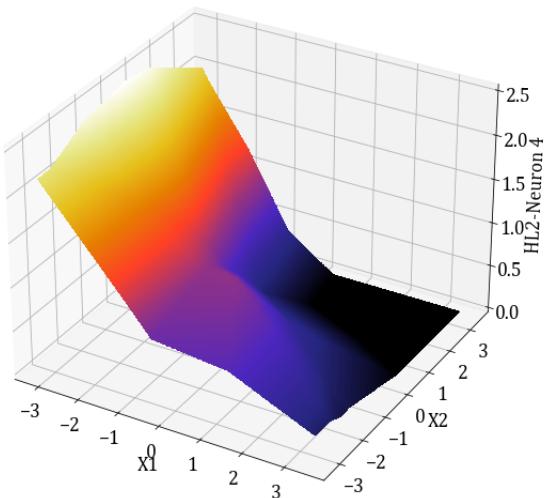
Epoch: 1; Surface for Layer 2, Neuron 4



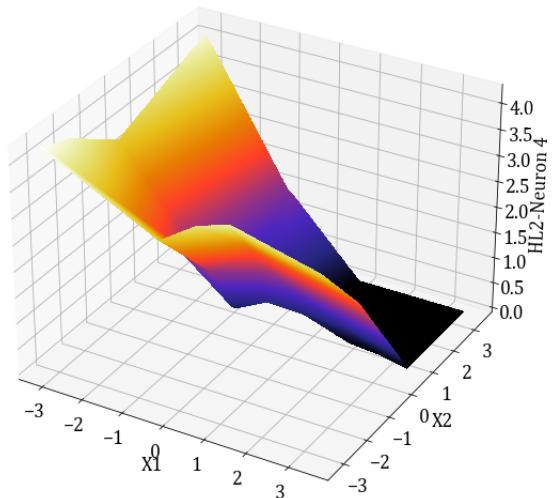
Epoch: 5; Surface for Layer 2, Neuron 4



Epoch: 20; Surface for Layer 2, Neuron 4



Epoch: 100; Surface for Layer 2, Neuron 4



Converged; Surface for Layer 2, Neuron 4

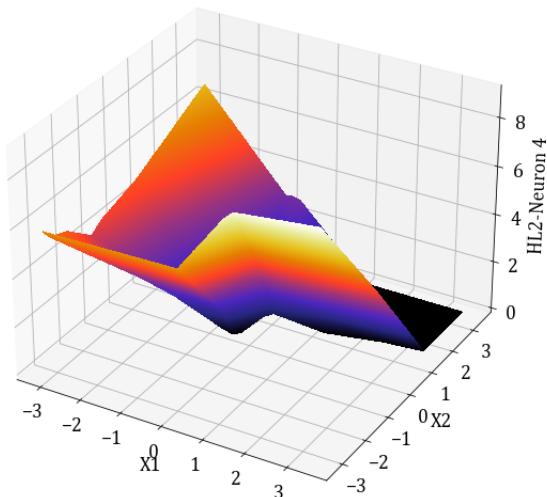
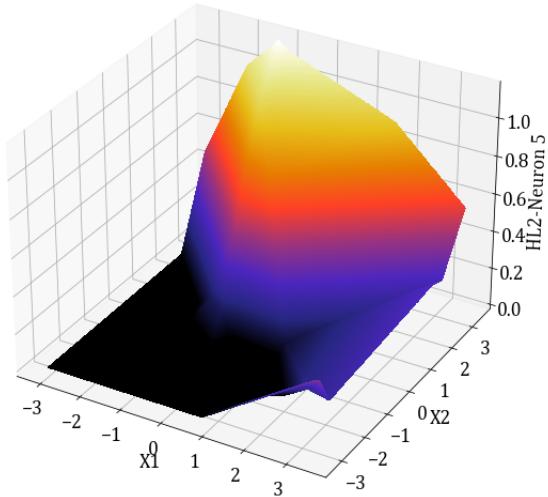


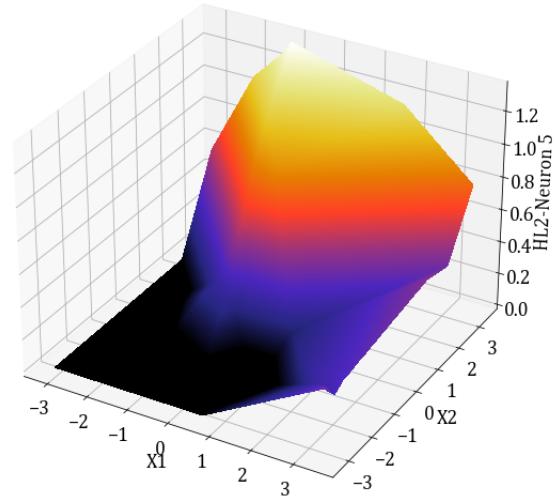
Figure 18: Surface Plots obtained for Hidden Layer 2, Neuron 4, across epochs.

2.1.4.13 Hidden Layer 2, Node 5

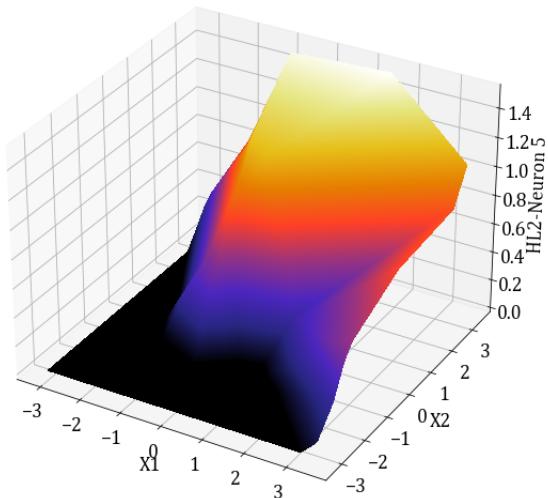
Epoch: 1; Surface for Layer 2, Neuron 5



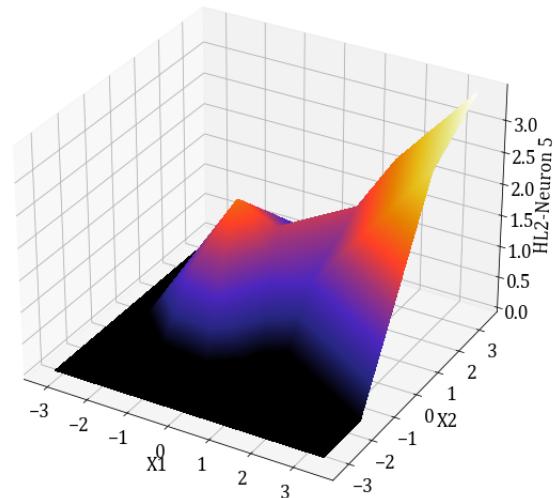
Epoch: 5; Surface for Layer 2, Neuron 5



Epoch: 20; Surface for Layer 2, Neuron 5



Epoch: 100; Surface for Layer 2, Neuron 5



Converged; Surface for Layer 2, Neuron 5

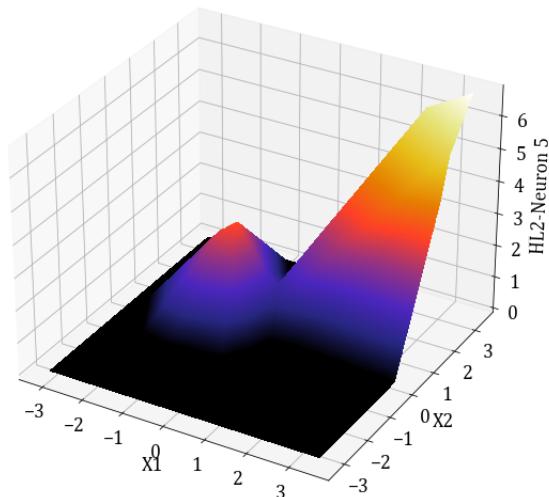
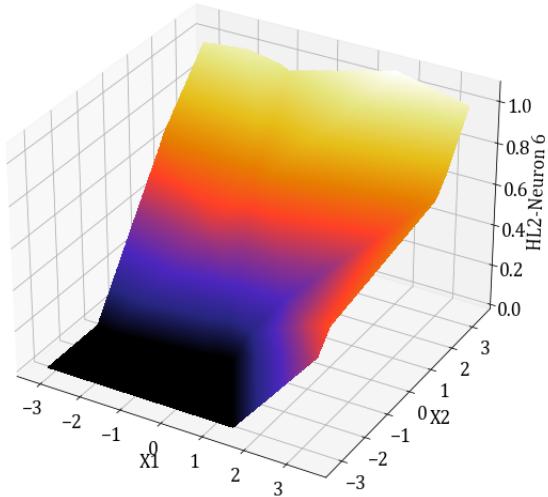


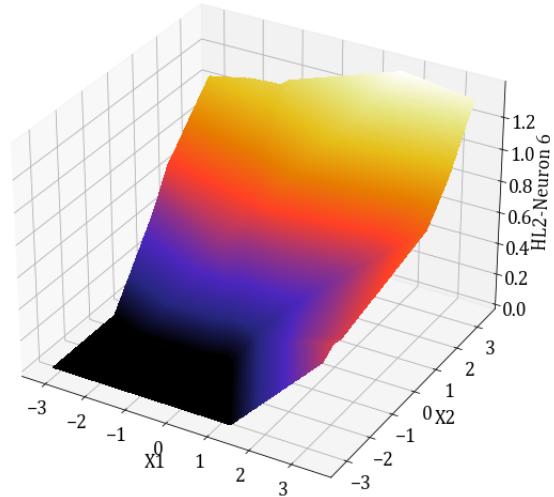
Figure 19: Surface Plots obtained for Hidden Layer 2, Neuron 5, across epochs.

2.1.4.14 Hidden Layer 2, Node 6

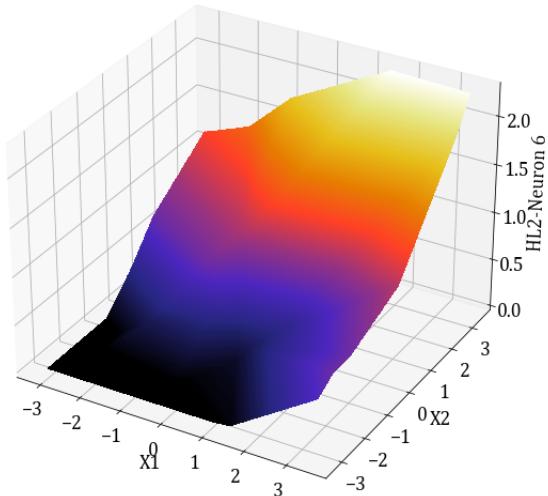
Epoch: 1; Surface for Layer 2, Neuron 6



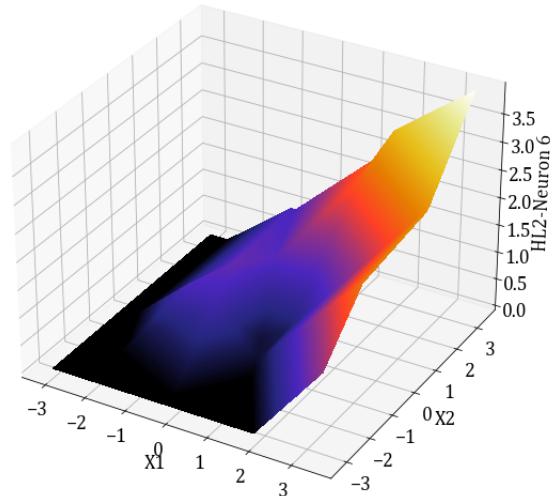
Epoch: 5; Surface for Layer 2, Neuron 6



Epoch: 20; Surface for Layer 2, Neuron 6



Epoch: 100; Surface for Layer 2, Neuron 6



Converged; Surface for Layer 2, Neuron 6

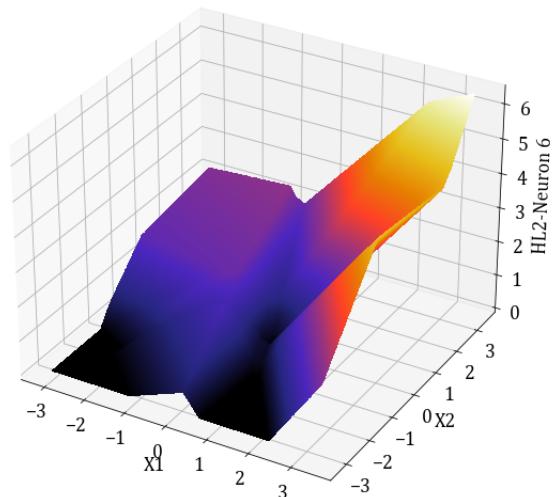
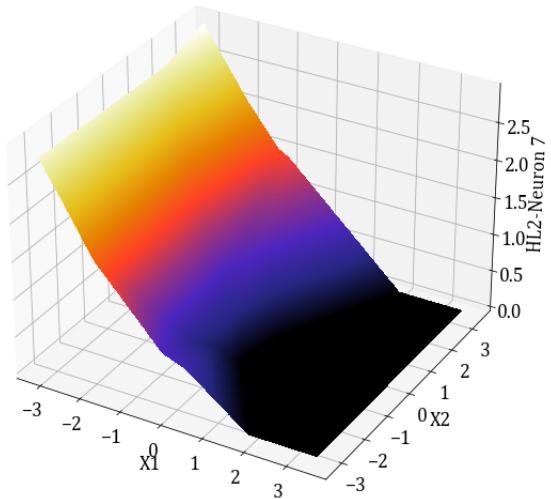


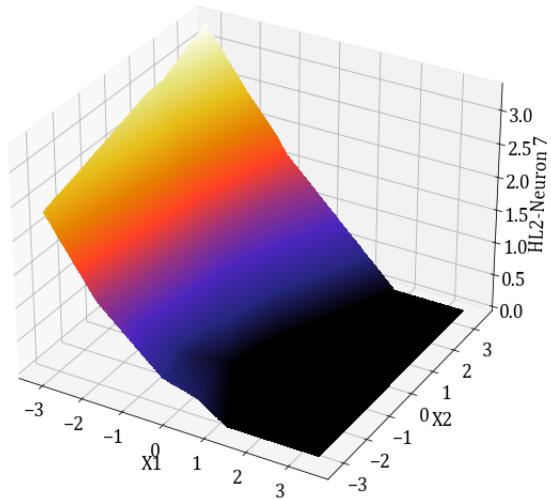
Figure 20: Surface Plots obtained for Hidden Layer 2, Neuron 6, across epochs.

2.1.4.15 Hidden Layer 2, Node 7

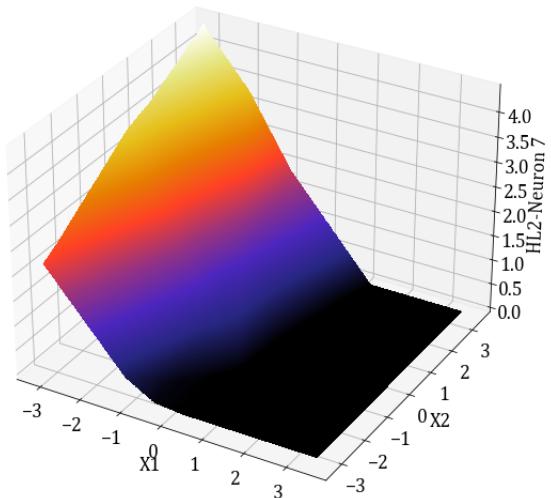
Epoch: 1; Surface for Layer 2, Neuron 7



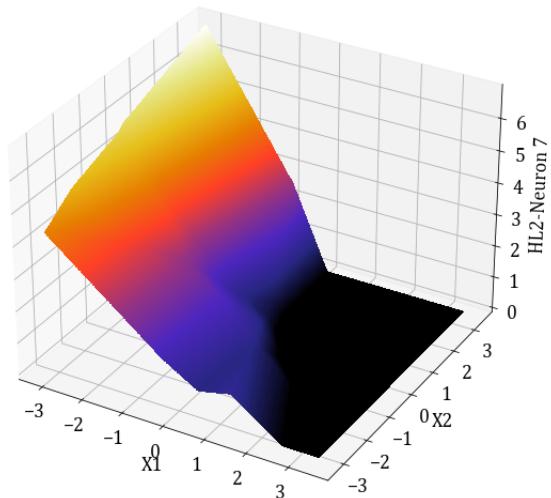
Epoch: 5; Surface for Layer 2, Neuron 7



Epoch: 20; Surface for Layer 2, Neuron 7



Epoch: 100; Surface for Layer 2, Neuron 7



Converged; Surface for Layer 2, Neuron 7

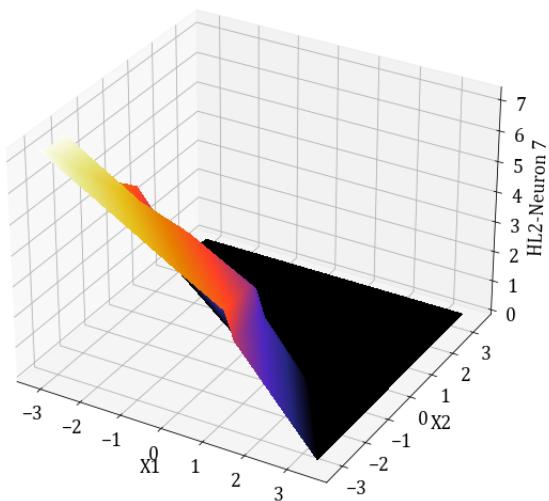
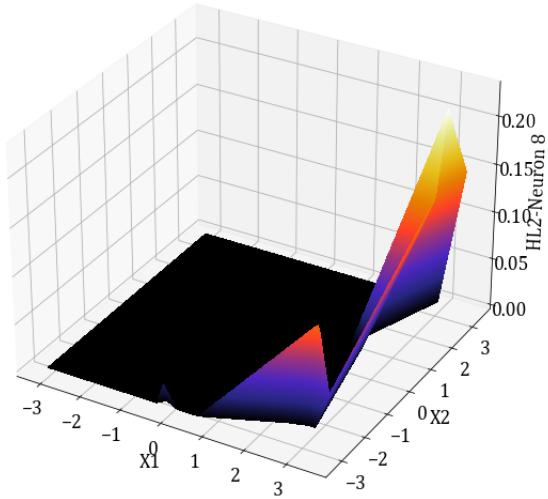


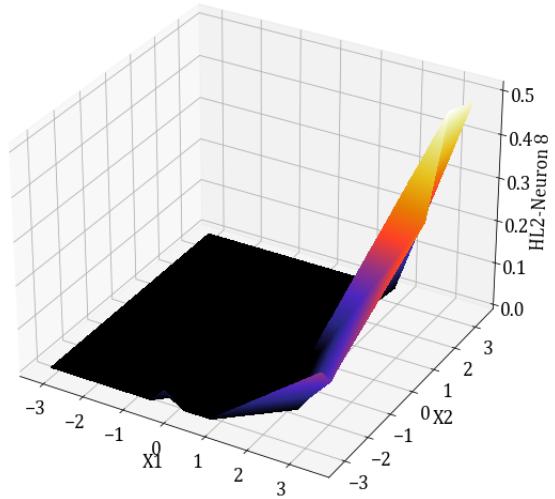
Figure 21: Surface Plots obtained for Hidden Layer 2, Neuron 7, across epochs.

2.1.4.16 Hidden Layer 2, Node 8

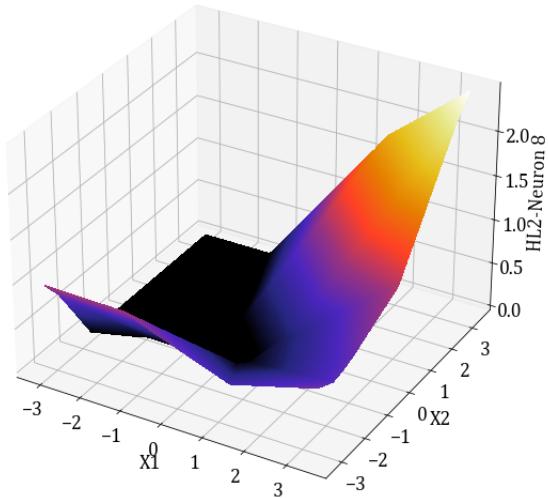
Epoch: 1; Surface for Layer 2, Neuron 8



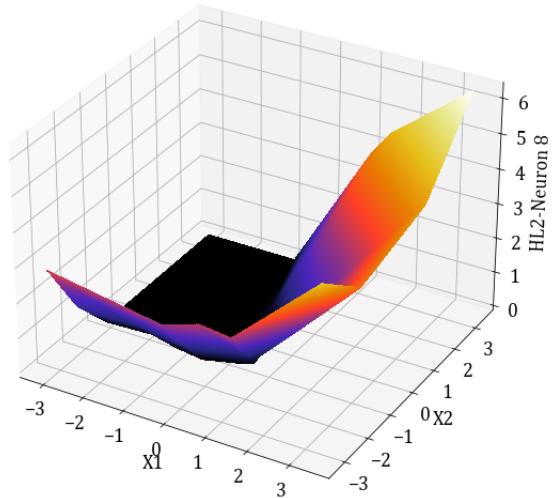
Epoch: 5; Surface for Layer 2, Neuron 8



Epoch: 20; Surface for Layer 2, Neuron 8



Epoch: 100; Surface for Layer 2, Neuron 8



Converged; Surface for Layer 2, Neuron 8

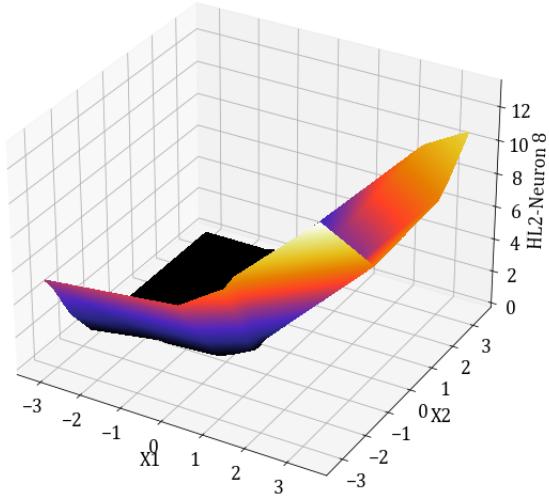
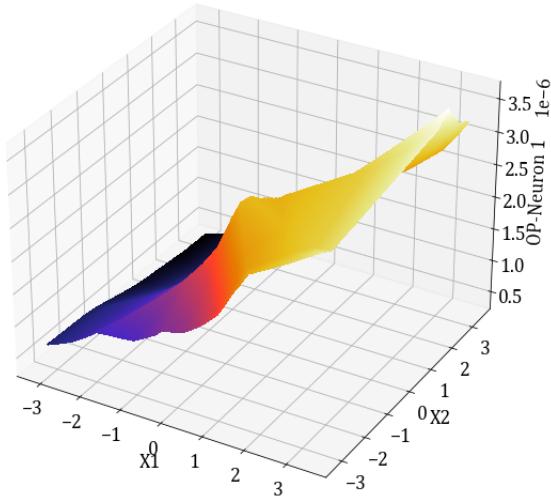


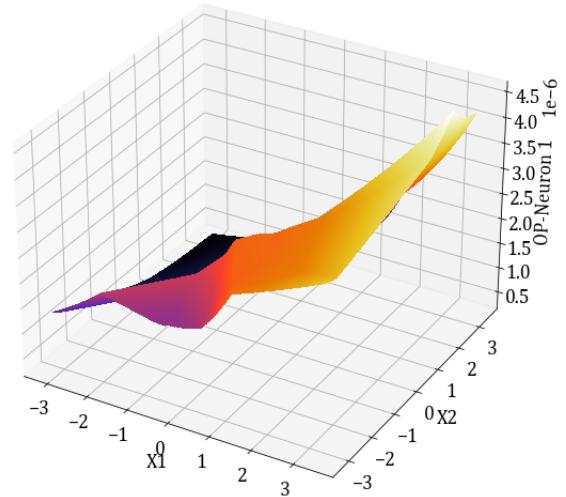
Figure 22: Surface Plots obtained for Hidden Layer 2, Neuron 8, across epochs.

2.1.4.17 Output Layer, Node 1

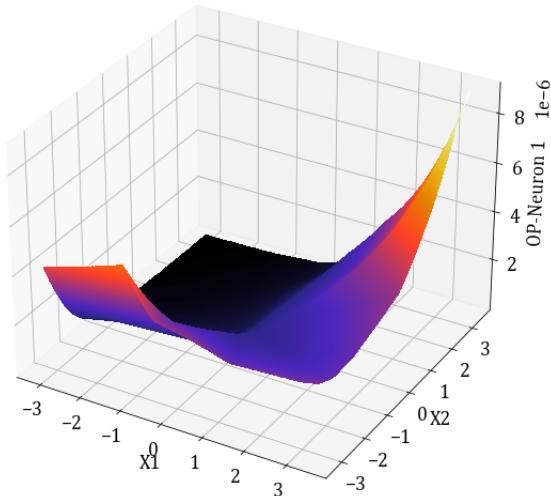
Epoch: 1; Surface for Output Layer, Neuron 1



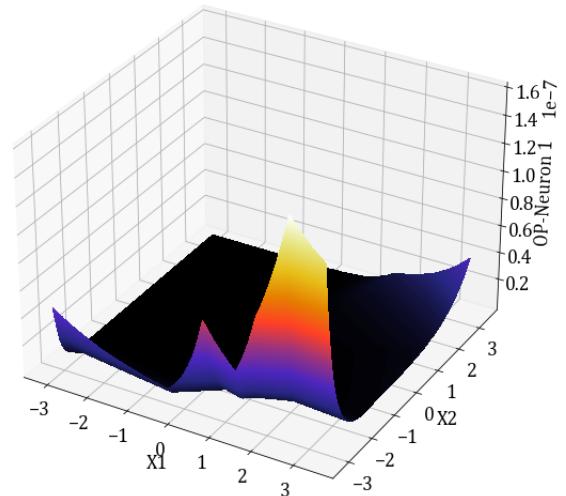
Epoch: 5; Surface for Output Layer, Neuron 1



Epoch: 20; Surface for Output Layer, Neuron 1



Epoch: 100; Surface for Output Layer, Neuron 1



Converged; Surface for Output Layer, Neuron 1

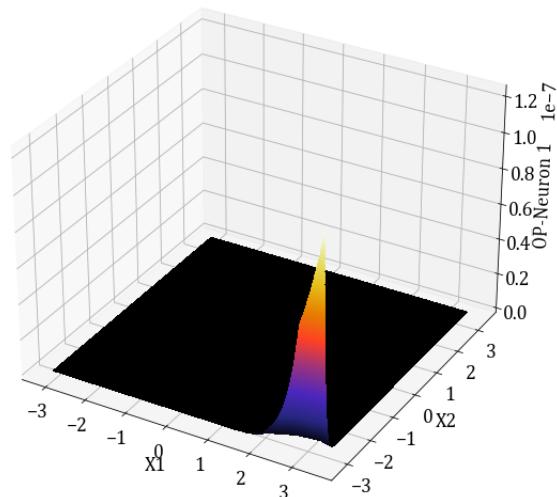
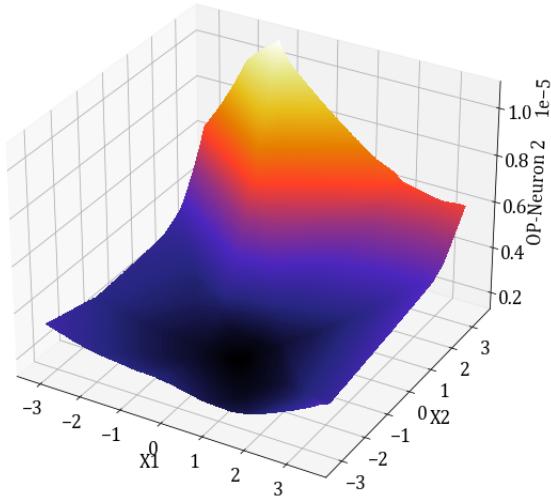


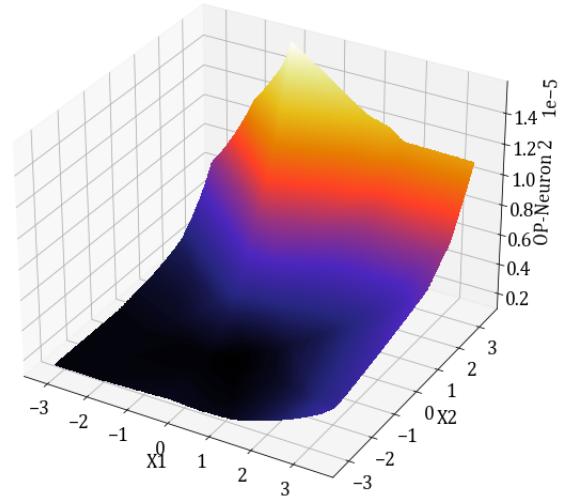
Figure 23: Surface Plots obtained for Output Layer, Neuron 1, across epochs.

2.1.4.18 Output Layer, Node 2

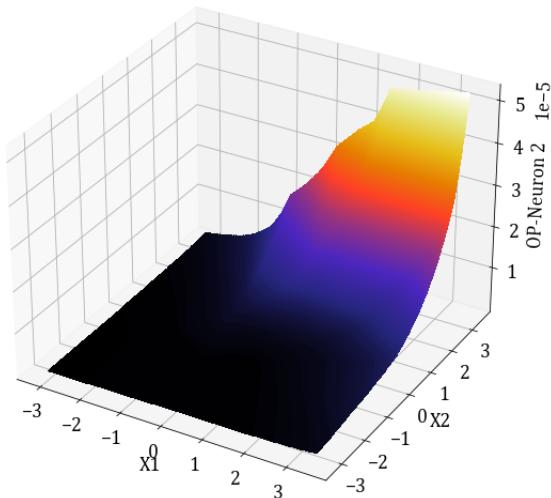
Epoch: 1; Surface for Output Layer, Neuron 2



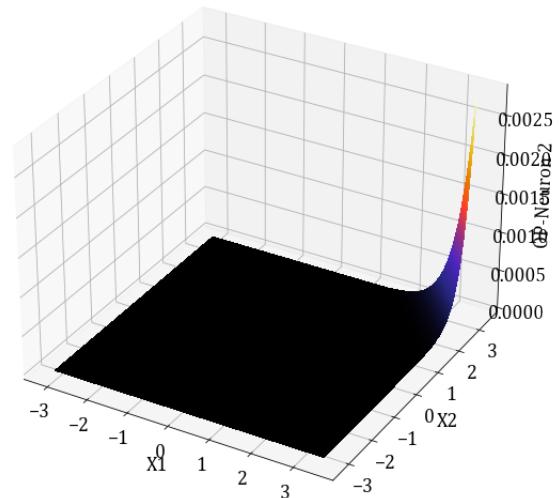
Epoch: 5; Surface for Output Layer, Neuron 2



Epoch: 20; Surface for Output Layer, Neuron 2



Epoch: 100; Surface for Output Layer, Neuron 2



Converged; Surface for Output Layer, Neuron 2

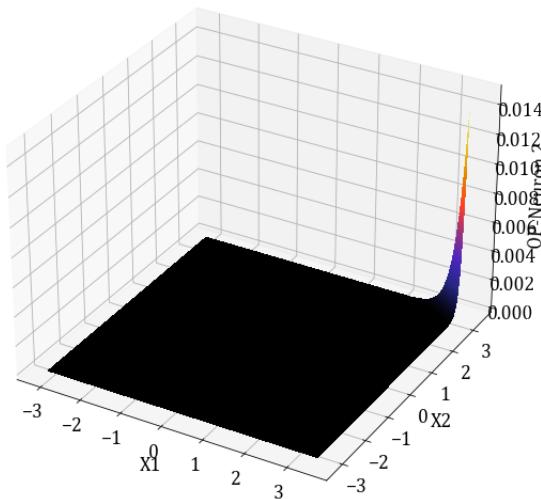
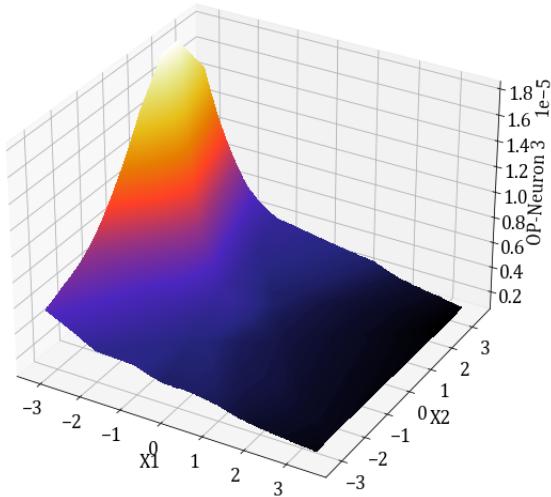


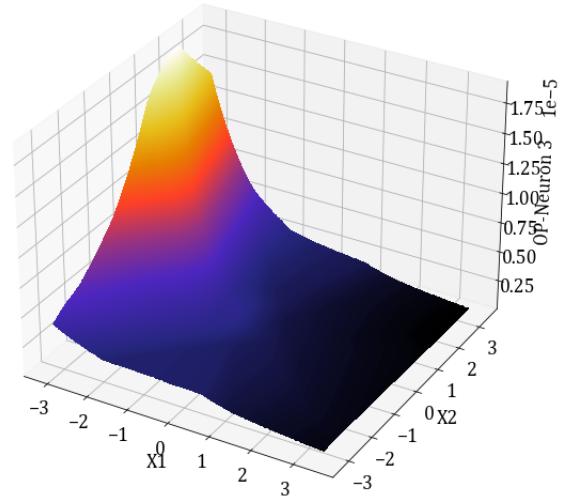
Figure 24: Surface Plots obtained for Output Layer, Neuron 2, across epochs.

2.1.4.19 Output Layer, Node 3

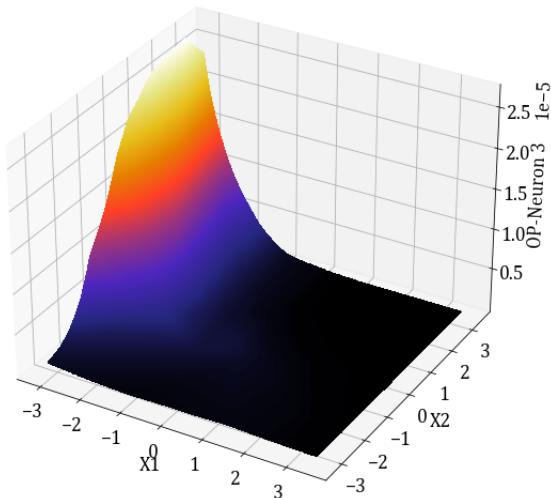
Epoch: 1; Surface for Output Layer, Neuron 3



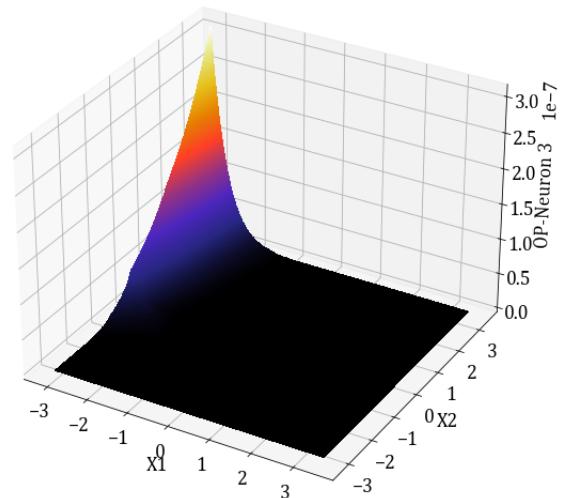
Epoch: 5; Surface for Output Layer, Neuron 3



Epoch: 20; Surface for Output Layer, Neuron 3



Epoch: 100; Surface for Output Layer, Neuron 3



Converged; Surface for Output Layer, Neuron 3

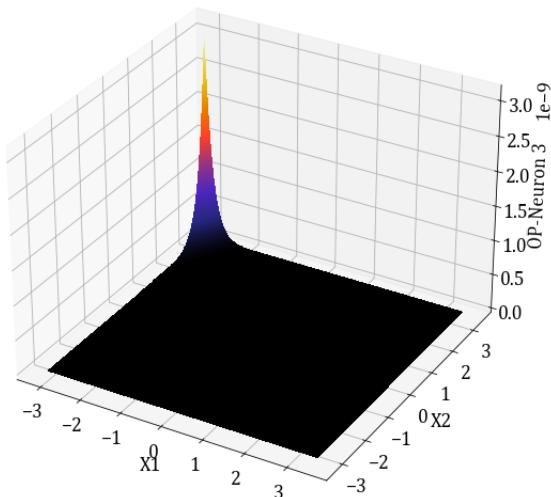


Figure 25: Surface Plots obtained for Output Layer, Neuron 3, across epochs.

From [Figure 7](#)-[Figure 25](#), we observe the following:

- First hidden layer surface plot is linear (activation function is [ReLU](#)).
- Non-Linearity is introduced in the surface plot of second hidden layer.
- The surface plot of the output neurons shows the selection cum localization of different classes in the latent space.

2.2 Non-Linear SVM

3 Dataset 2A

3.1 MLFFNN

3.2 Gaussian-kernel SVM