

CS7015 : DEEP LEARNING

PROGRAMMING ASSIGNMENT

Backpropagation

February 17, 2018

Student ID: EE15B123 & EE15B025

Namida M

Ganga Meghanath

Contents

1	Introduction	2
2	Supported Hyperparameters	2
3	Plots	2
3.1	No. of Hidden Layers : 1	2
3.2	No. of Hidden Layers : 2	4
3.3	No. of Hidden Layers : 3	6
3.4	No. of Hidden Layers : 4	8
3.5	Adam, NAG, Momentum, GD	10
3.6	sigmoid v/s tanh	12
3.7	cross entropy loss v/s squared error loss	14
4	Best Model	16

1 INTRODUCTION

This is a basic implementation of neural network back-propagation algorithm using python. Comparison between the functionality of the neural network for different sets of hyperparameters have been demonstrated.

2 SUPPORTED HYPERPARAMETERS

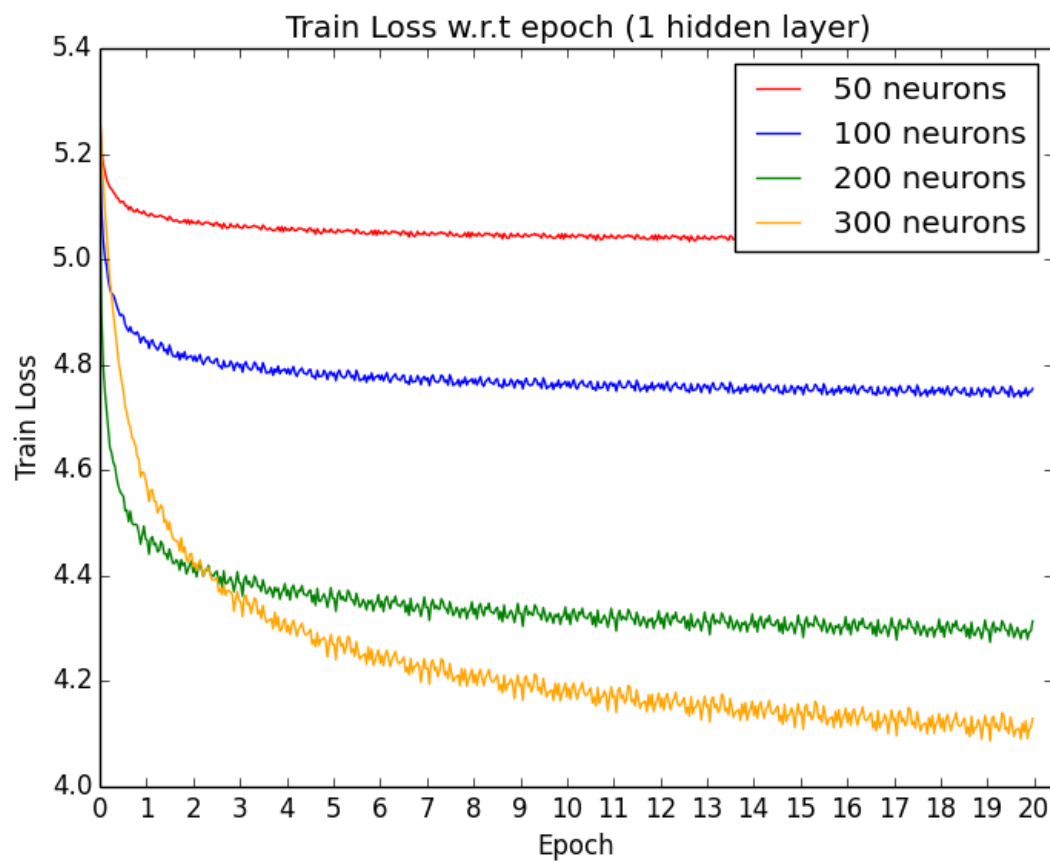
Hyperparameters	
Activation	sigmoid, tanh
Loss	Cross Entropy, Squared Error
Optimiser	Adam, GD, Momentum, NAG
Anneal	True, Falses

3 PLOTS

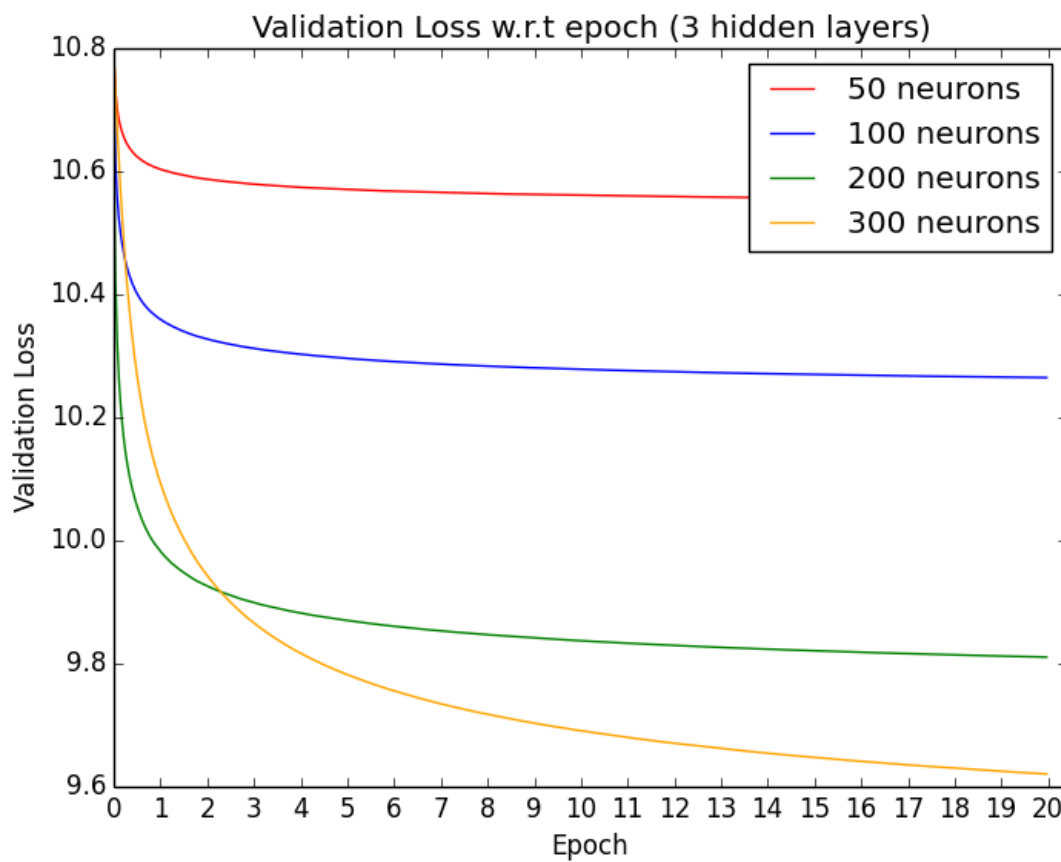
3.1 No. of Hidden Layers : 1

Hyperparameters used :

Hyperparameters	
Activation	Sigmoid
Loss	Cross Entropy
Optimiser	Adam
Batch size	20



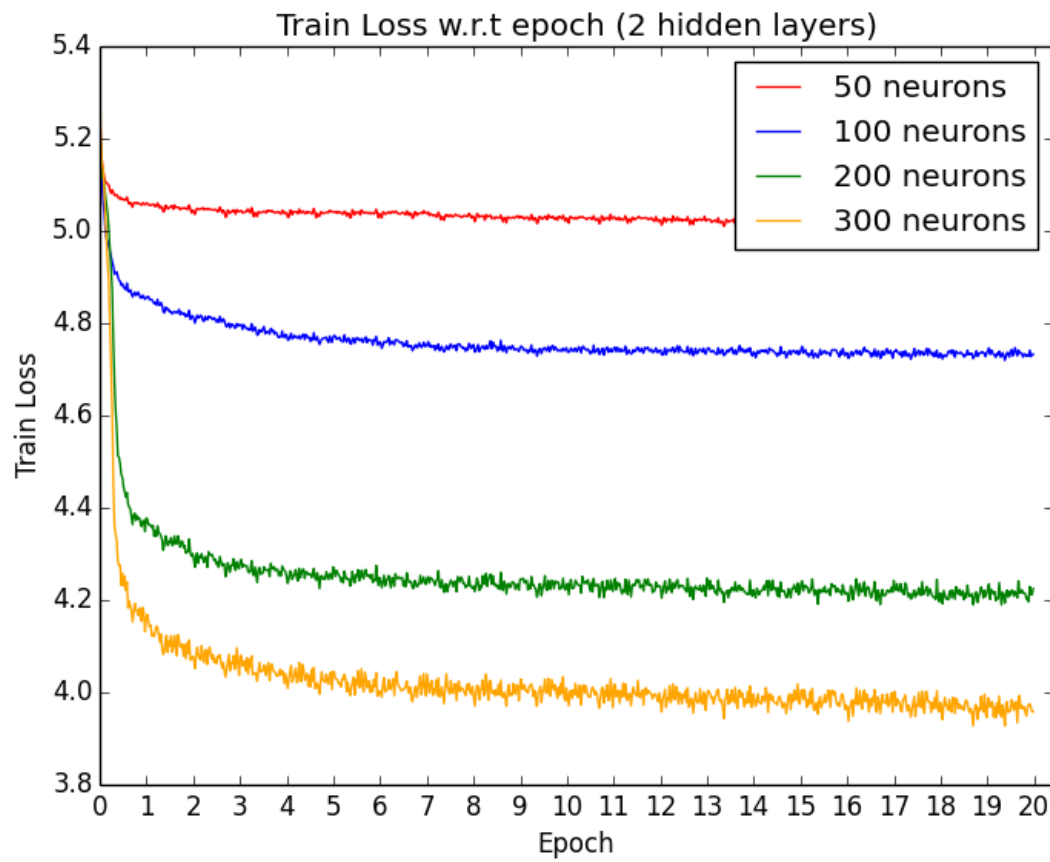
No. of Neurons in Hidden Layer	Validation accuracy(%) for 20 epochs
50	75.92
100	81
200	81.02
300	76.66



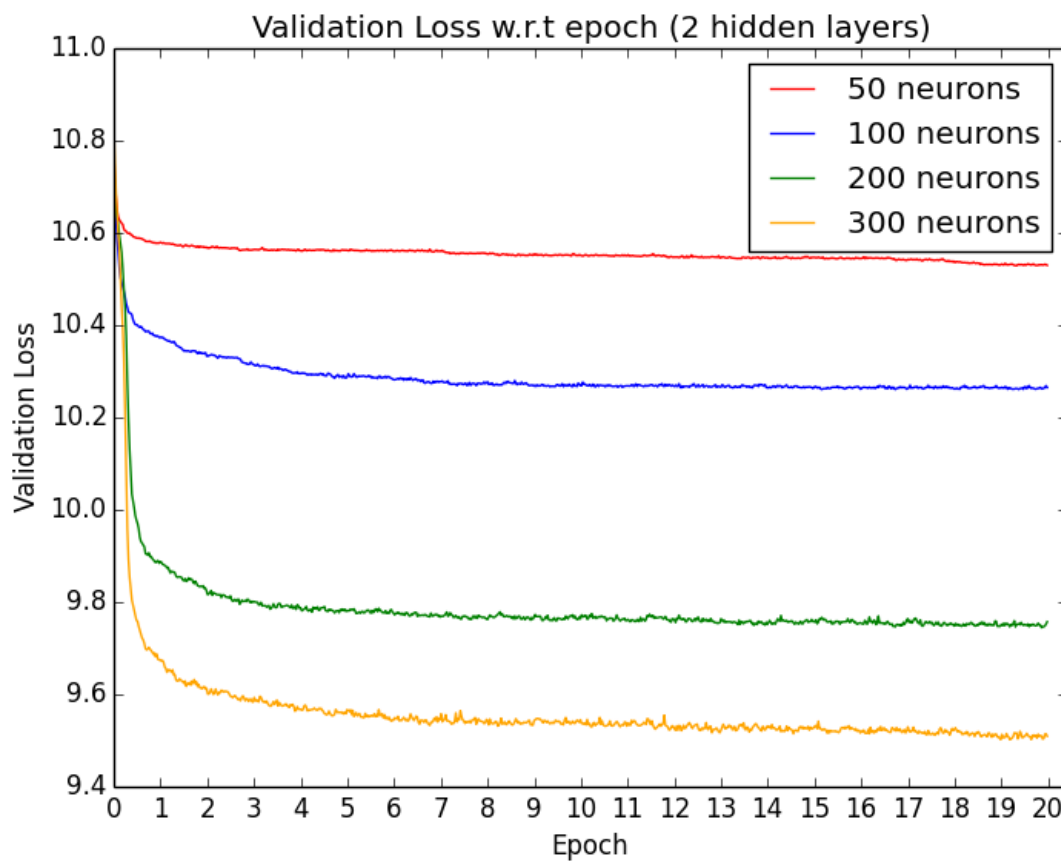
3.2 No. of Hidden Layers : 2

Hyperparameters used :

Hyperparameters	
Activation	Sigmoid
Loss	Cross Entropy
Optimiser	Adam
Batch size	20



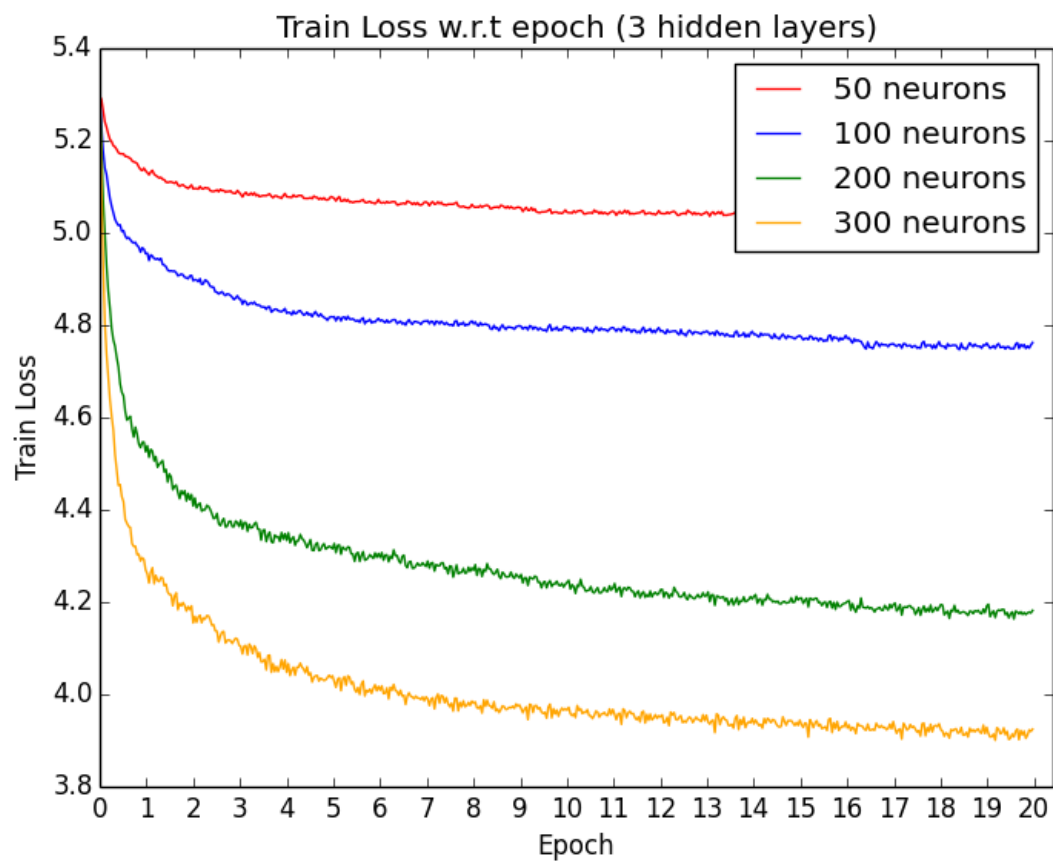
No. of Neurons in Hidden Layer	Validation accuracy(%) for 20 epochs
50	79.9
100	81.5
200	85.26
300	86.5



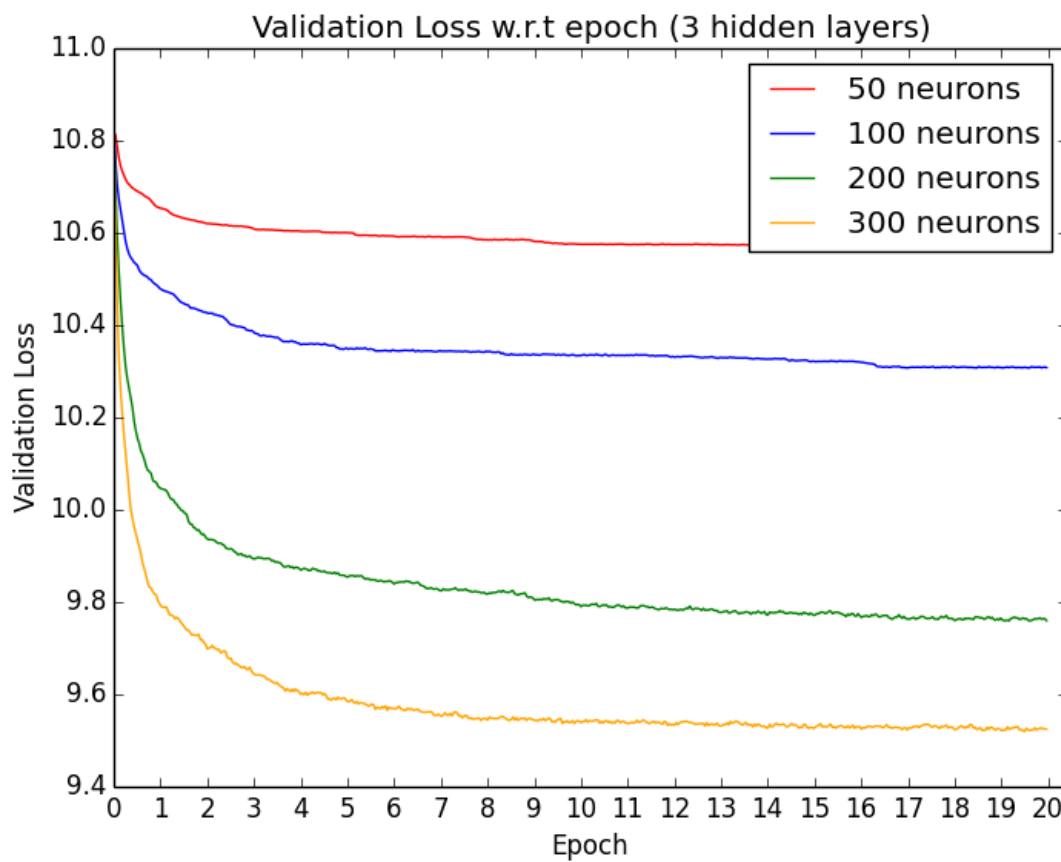
3.3 No. of Hidden Layers : 3

Hyperparameters used :

Hyperparameters	
Activation	Sigmoid
Loss	Cross Entropy
Optimiser	Adam
Batch size	20



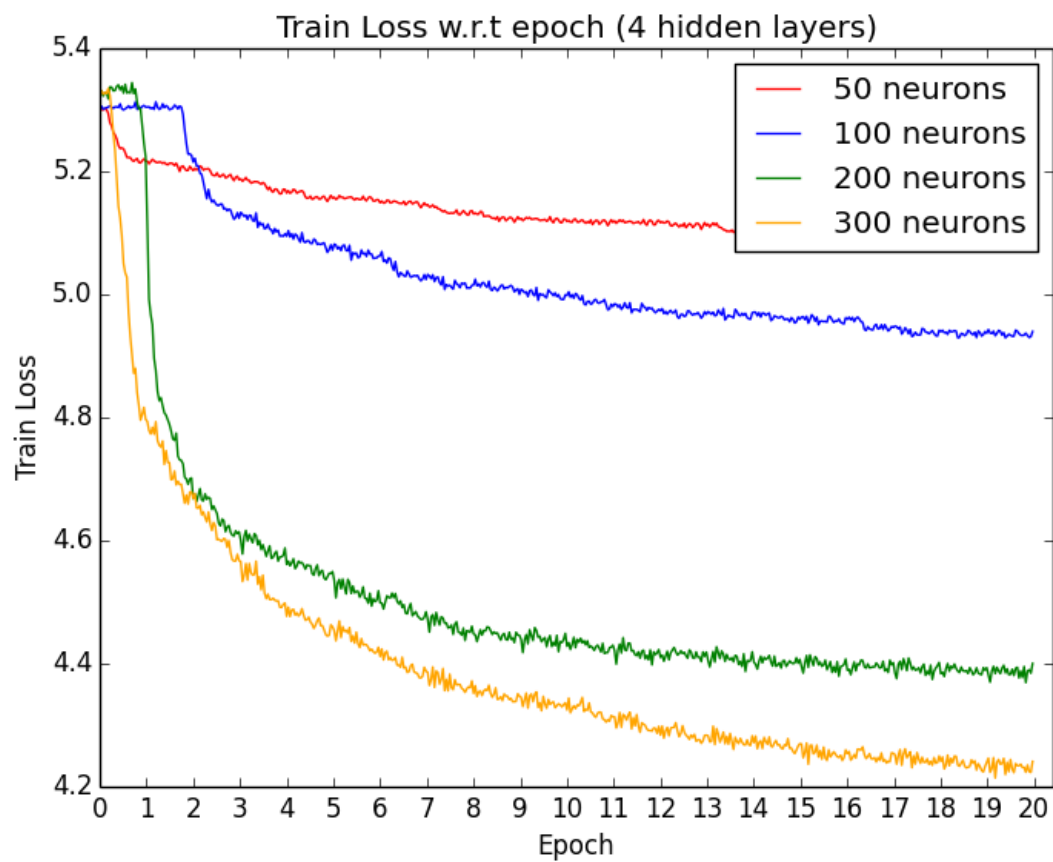
No. of Neurons in Hidden Layer	Validation accuracy(%) for 20 epochs
50	86.54
100	87.64
200	88.26
300	87.82



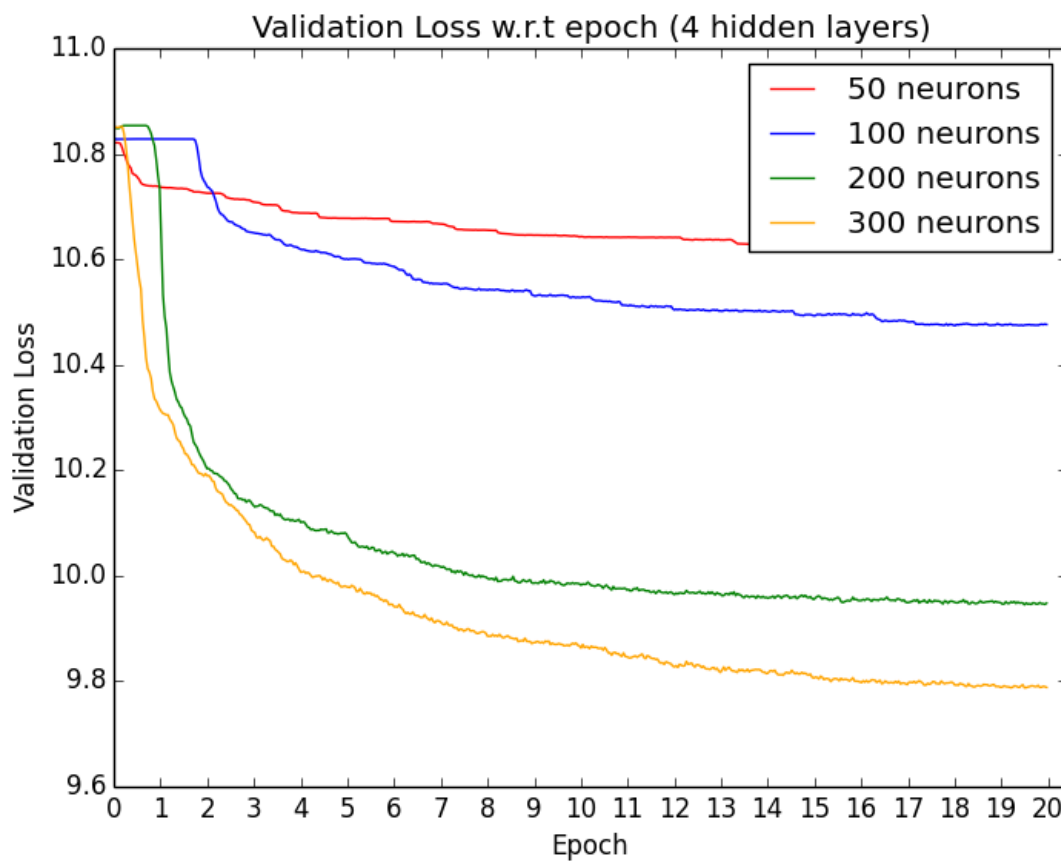
3.4 No. of Hidden Layers : 4

Hyperparameters used :

Hyperparameters	
Activation	Sigmoid
Loss	Cross Entropy
Optimiser	Adam
Batch size	20



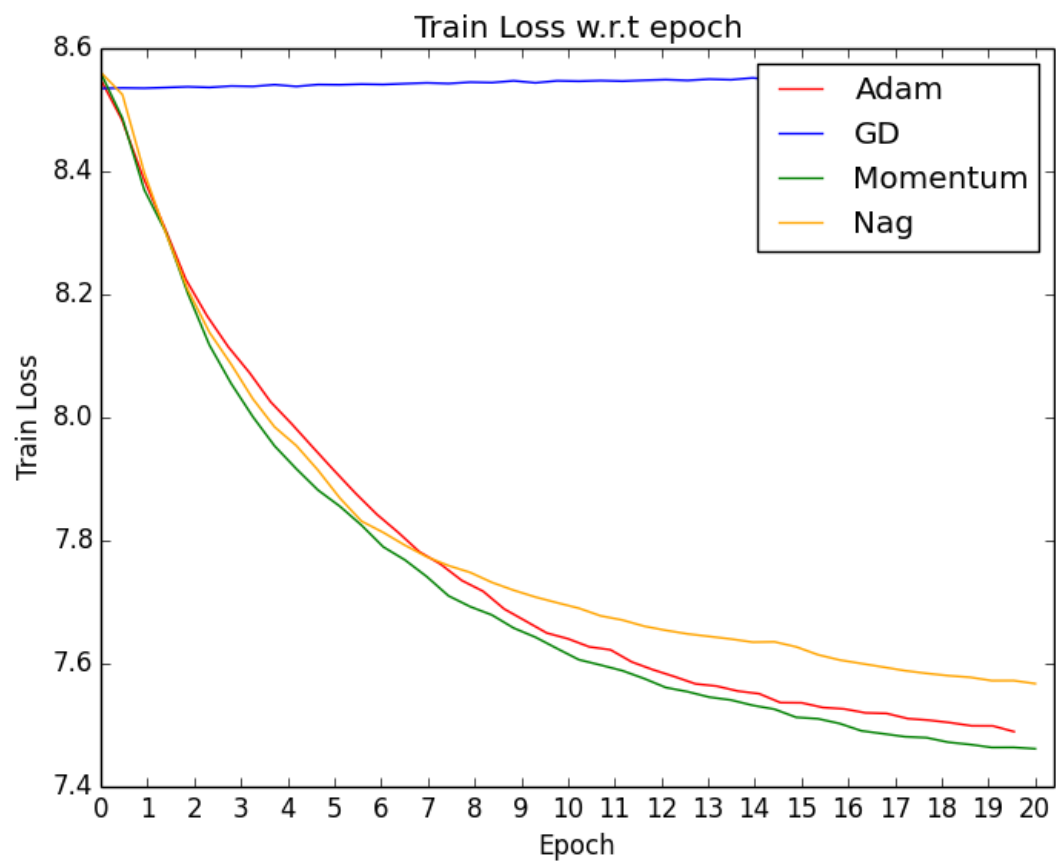
No. of Neurons in Hidden Layer	Validation accuracy(%) for 20 epochs
50	64.6
100	77.76
200	87.62
300	87.2

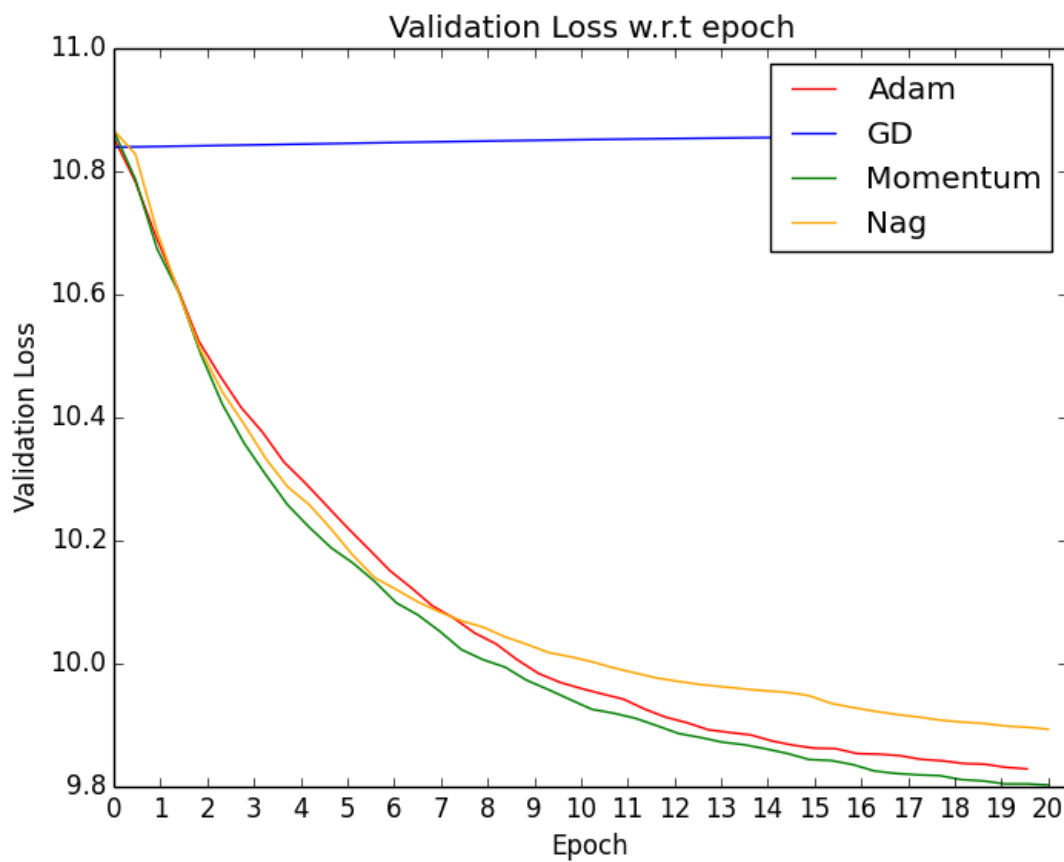


3.5 Adam, NAG, Momentum, GD

Hyperparameters used :

Hyperparameters	
Activation	Sigmoid
Loss	Cross Entropy
No. of Hidden Layers	3 (300,300,300)
Batch size	20





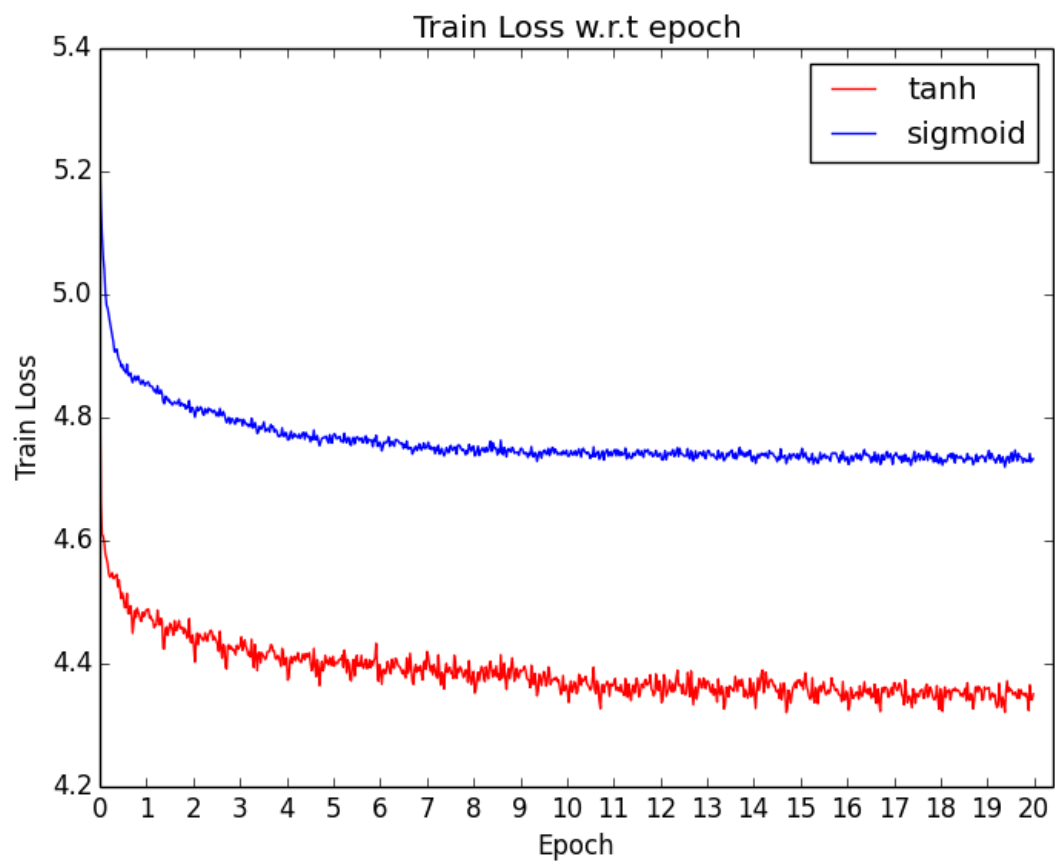
Adam :

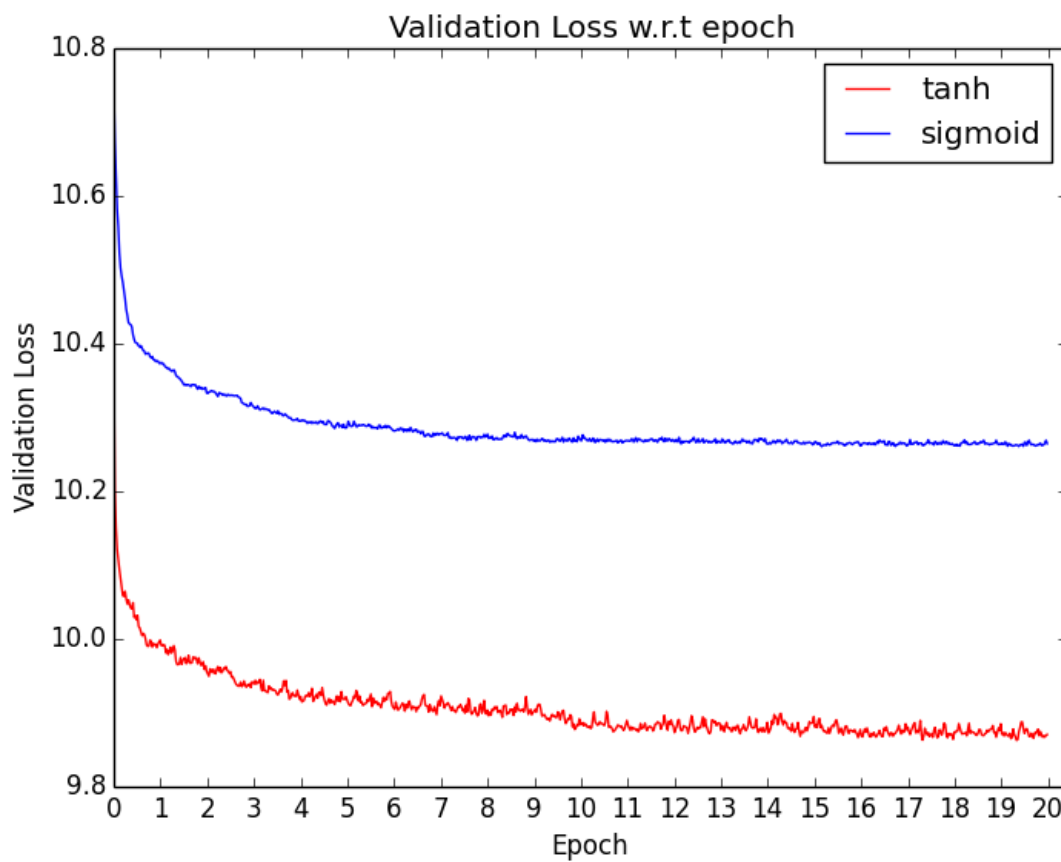
Nag :

3.6 sigmoid v/s tanh

Hyperparameters used :

Hyperparameters	
Optimiser	Adam
Loss	Cross Entropy
No. of Hidden Layers	2 (100,100)
Batch size	20

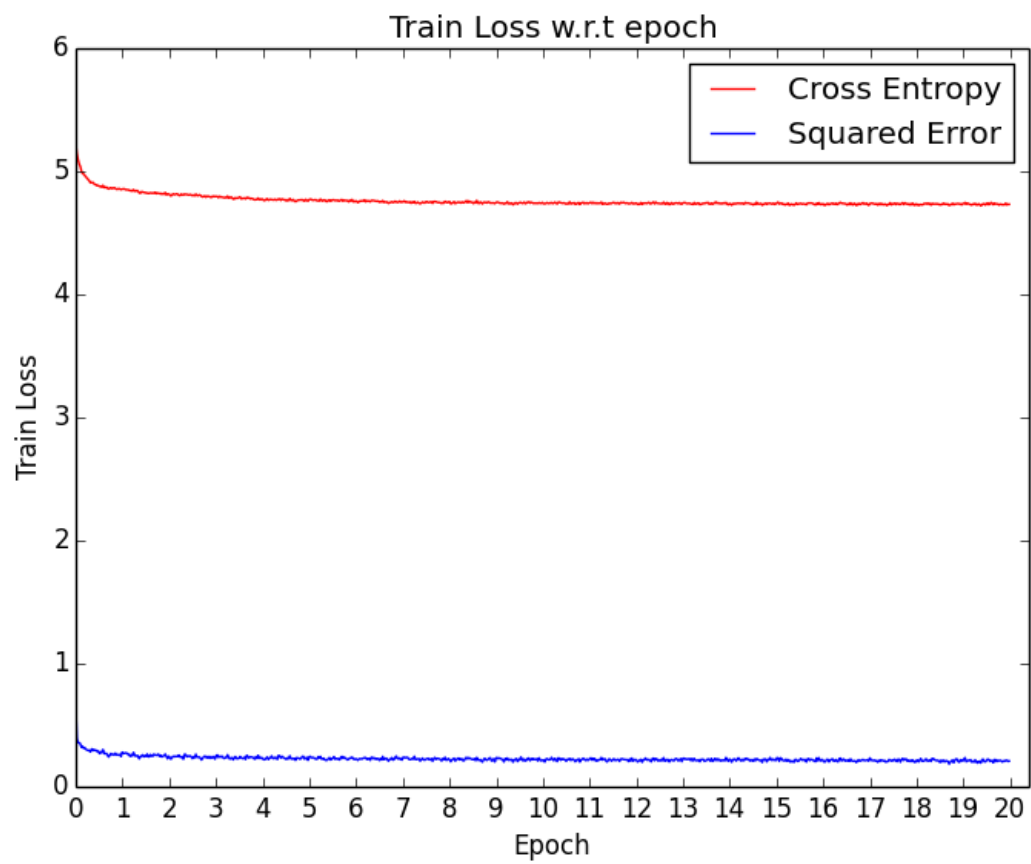


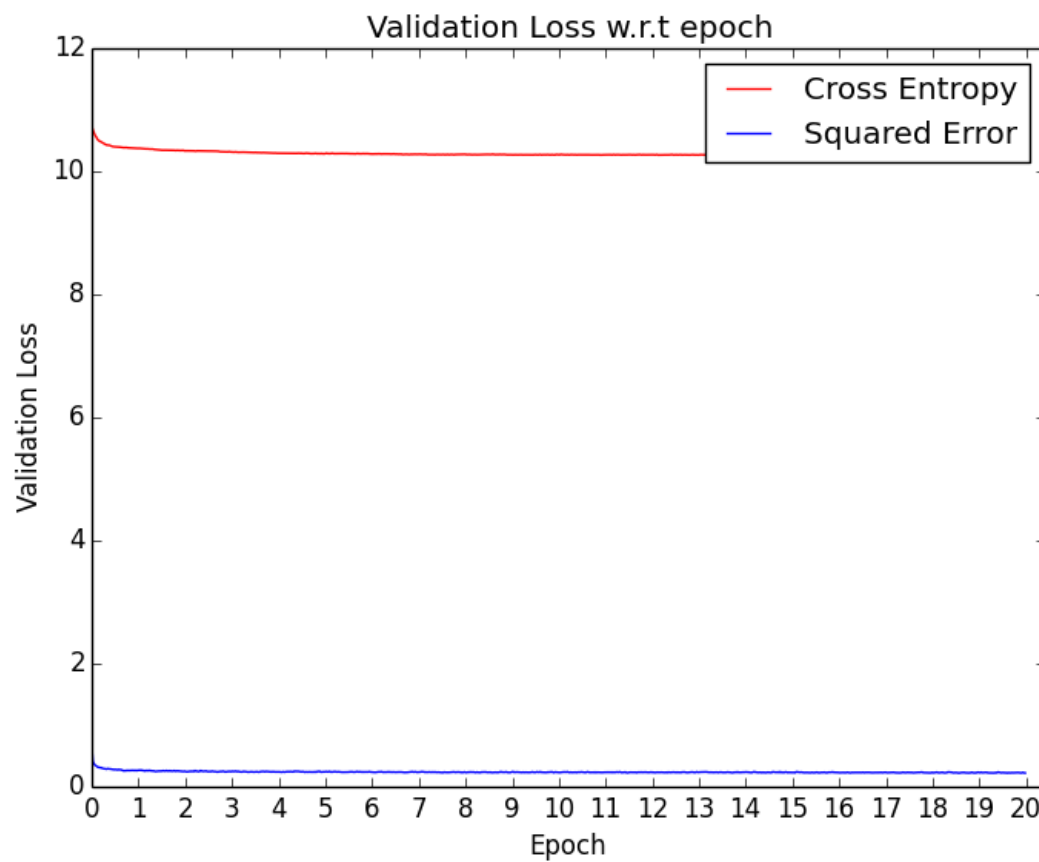


3.7 cross entropy loss v/s squared error loss

Hyperparameters used :

Hyperparameters	
Activation	Sigmoid
Optimiser	Adam
No. of Hidden Layers	2 (100,100)
Batch size	20





4 BEST MODEL

Hyperparameters used :

Hyperparameters	
Activation	tanh
Optimiser	Adam
No. of Hidden Layers	3 (128,256,256)
Batch size	20
Loss	Cross Entropy
Learning rate	0.001

Kaggle score = 88.766

Accuracy on validation = 88.6

Other Submissions :

1. Submission filename : test_submission.csv

Hyperparameters used :

Hyperparameters	
Activation	tanh
Optimiser	Adam
No. of Hidden Layers	1 (300)
Batch size	20
Loss	Cross Entropy
Learning rate	0.001

Kaggle score = 82.066

Accuracy on validation = 82

2. Submission filename : test_submission100.csv

Hyperparameters used :

Hyperparameters	
Activation	tanh
Optimiser	Adam
No. of Hidden Layers	3 (100,100,100)
Batch size	20
Loss	Cross Entropy
Learning rate	0.001

Kaggle score = 87.06

Accuracy on validation = 88.26

3. Submission filename : test_submission300.csv

Hyperparameters used :

Hyperparameters	
Activation	tanh
Optimiser	Adam
No. of Hidden Layers	3 (300,300,300)
Batch size	20
Loss	Cross Entropy
Learning rate	0.001

Kaggle score = 88.4

Accuracy on validation = 88.26

4. Submission filename : test_submission300.csv

Hyperparameters used :

Hyperparameters	
Activation	tanh
Optimiser	Adam
No. of Hidden Layers	3 (300,300,300)
Batch size	20
Loss	Cross Entropy
Learning rate	0.001

Kaggle score = 88.4

Accuracy on validation = 88.26

5. Submission filename: sub_2.csv

Hyperparameters used :

Hyperparameters	
Activation	sigmoid
Optimiser	Adam
No. of Hidden Layers	3 (300,300)
Batch size	20
Loss	Cross Entropy
Learning rate	0.001

Kaggle score = 86.5

6. Submission filename: sub3.csv Hyperparameters used :

Hyperparameters	
Activation	sigmoid
Optimiser	Adam
No. of Hidden Layers	2 (200,200)
Batch size	200
Loss	Cross Entropy
Learning rate	0.005

Kaggle score = 88.4

NOTE :

- Even though the best **kaggle** submission was observed for the above set of hyperparameters, we got to know we had to pickle the model later on and hence due to the unavailability of resources and time(as other codes were running during the time), the weights submitted are for a sub-optimal model which gives an accuracy around 0.5% less than our best submission.
- **sklearn** has been used only for calculating accuracy and for no other purpose

Bibliography

- [1] Mitesh M Khapra. *CS7015 Deep Learning: Lecture 5*, Indian Institute of Technology Madras, 2018