# ASSIGNMENT 3: AI42001:Fashion MNIST

Submitted by:

- 21EE30038- Chinmay Gupta
- 21EE30039- Ajay Paliwal
- 21ME10123- Abhinav R Cillanki
- 19MF3AI23- Sachin
- 21EE30012- Harsh Singh Chauhan
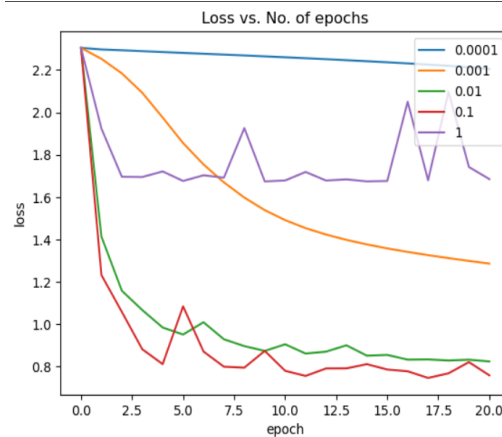- 21EE10073- Sushupthi R Koushik

QUESTIONS:

Q3 AND Q5. Training models with corresponding Learning Rates with SGD (momentum = 0.9) and <u>Mini-Batch size = 64</u> for <u>20 epochs</u> and plotting the Training Loss Vs. Epoch number.
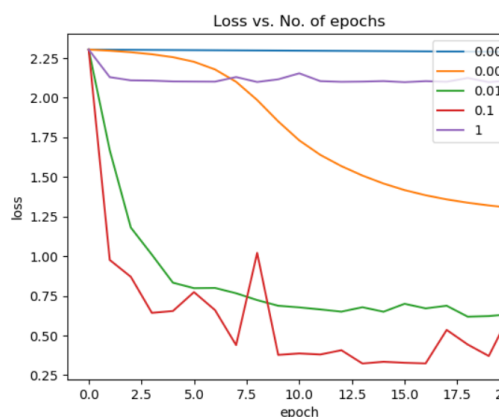
MODEL 1
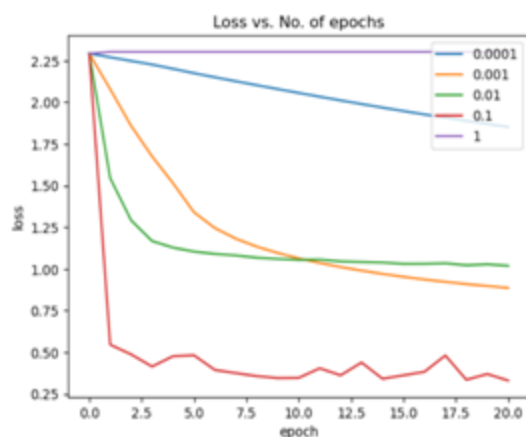
MODEL 2

MODEL 3

Q4 AND Q6.

The learning rate of SGD was fixed at 0.01 and different mini-batch sizes were tried, again for <u>20 epochs</u>. After that the models were tested on the test set and the test set accuracies was noted.

Models	Batch Size 1	Batch Size 2	Batch Size 3	Batch Size 4	Batch Size 5
Model 1	87.42	88.35	88.67	88.68	88.71
Model 2	87.59	88.76	89.25	89.49	89.32
Model 3	88.14	87.80	89.16	89.38	89.26

Q7.

Finally, the model in 2(a) was swapped with the model 784-512-10. Using learning rate 0.01 with batch size of 64, this model was trained and tested on the test set. The corresponding accuracies under same batch size and learning rate are as follows.
The Model: 784-512-10 reported a test set accuracy of: **84.49%**.

The Model: 784-256-10 reported a test set accuracy of: **64.72%.**

Q8) OBSERVATIONS:
a) How does increasing the learning rate affect the training loss?

From the three plots attached above, it is observed that the training loss is very high for the very small learning rates and decreases with increase in learning rate but after a certain value, the loss function starts to increase with increase in learning rate. Example: Loss decreases as learning rate increases from 0.001 up to 0.01. But from there onwards loss increases as learning rate increases from 0.01 to 1.

b) How does increasing the batch size affect validation accuracy?

As observed from the above table, the increase in batch size initially improves the accuracy and on further increase, there is a decrease in accuracy.

c) How does increasing depth affect validation accuracy?

From the above table, we can observe that the accuracy generally increases with the increase in depth of the Neural Network.

d)How does increasing the number of parameters affect validation accuracy?

As evident from the table above, the validation accuracy generally increases with the increase in number of learnable parameters.

Q9) REASONS:

a) How does increasing the learning rate affect the training loss?

When we use gradient descent for parameter updates, very low learning rate slows down the process and it takes a lot of time to reach to the global minima of the loss function, so the loss function remains high after each epoch. Meanwhile, with high learning rates, we tend to overshoot the minima while learning, which again increases the loss.

b) How does increasing the batch size affect validation accuracy?

While performing gradient descent using mini-batches, we, in a way, take up the average of the gradients of the entire mini-batch. This way, larger mini-batches provide a stability to the gradients and prevents overshooting the minima. But if the mini-batch size is too large, it reduces the stochasticity of the gradients and hence have a negative effect on the accuracy.

c) How does increasing depth affect validation accuracy?

With the increasing depth and complexity of the Neural Network, we allow the model to learn more complex features which generally can't be done with shallow networks. Thus, the validation accuracy increases with increasing the depth of the neural network.

d) How does increasing the number of parameters affect validation accuracy?

The number of learnable parameters directly corresponds to increase in depth and/or width of the network. The deeper the network is, the more complex features it can learn. Hence, with increase in complexity of the network, the validation accuracy generally tends to increase.