

Week 6 Assignment:

Training LeNet on Cifar 10 dataset with different optimizer.

Solution:

After Training the LeNet Model with all optimizer in provided link , the results of the training are shown in the table below:

The model was trained with only 10 epochs for each optimizer.

Optimizer	Convergence rate (loss)	Accuracy	Average Processing time for each epoch
MomentumOptimizer	2.22	0.15	274.5 ms
GradientDescentOptimizer	1.09	0.65	310.78 ms
AdadeltaOptimizer	1.92	0.32	283.06
AdagradOptimizer	1.19	0.53	283.85 ms
AdagradDAOptimizer	2.01	0.19	296.11 ms
AdamOptimizer	1.02	0.61	281.8 ms
FtrlOptimizer	1.66	0.37	281.29 ms
ProximalGradientDescentOptimizer	1.08	0.66	283.38 ms
ProximalAdagradOptimizer	1.29	0.54	280.88 ms
RMSPropOptimizer	0.89	0.68	279.43 ms

From the table above, we can notice the effects of optimizers on our model and its learning performance, beside many hyperparameters that we could tune to improve the performance of the neural network the choice of the optimizer is what makes the biggest difference. Accordingly, we can notice that the RMSPropOptimizer gives a higher accuracy with less processing time comparing to the other optimizers. The RMSPropOptimizer has a simple and effective method to find the optimum values for the neural network. The RMSprop optimizer restricts the oscillations in the vertical direction. Therefore, we can increase our learning rate and our algorithm could take larger steps in the horizontal direction converging faster. In conclusion, Optimizers are one of the vital parts in neural networks understanding how they work would help in improving the way our models perform