COL-780
ASSIGNMENT-4 REPORT

BY

KURISETI RAVI SRI TEJA

2019CS10369

Contents

1 Experiments Done

1.1 Optimizers

- I tried with optimizers such as Adam, AdamW, and SGD with momentum.
- Among all the three optimizers, SGD with momentum = 0.9 gave the best results and faster convergence.

1.2 Learning Rates

- I experimented with learning rates of 1e-1, 1e-2, 1e-3, 1e-4, and 1e-5.
- With learning rates of 1e-1 and 1e-2, the model did not converge at all.
- With learning rates of 1e-3, 1e-4, and 1e-5, the model converged in 40 epochs, and the validation accuracy for the best epochs was around 95–99
- The learning rate of 1e-3 was best because the model converged within 20 epochs.

1.3 Weight Decay

All experiments were done with a learning rate of 0.001 for 20 epochs using SGD with momentum = 0.9 as the optimizer with L2 regularization.

1.3.1 Results

Weight Decay Parameter	Training Loss	Best Validation Accuracy	Test Accuracy on Best Model
1	3.218	4%	4%
0.1	2.858	48%	46.4%
0.01	0.326	94.4%	71.2%
0.001	0.001	100%	99.2%
0.00001	0.001	98.4%	98.4%
0.000001	0.001	99.2%	98.4%

- We see that with a weight decay of 1e-3, the model gives the best results.
- For high values of weight decay, the model does not learn anything and hence the performance is very low.

2 Instructions to run the code

Use the command **python3 Code.py arg1** to run the code where arg1 is the path to the root directory containing train,valid and test directories. On executing the code runs for 20 epochs and saves the weights of the epoch with the best validation accuracy.