Backpropagation Algorithm Assignment

Jay Jagtap, North Carolina State University

I. PROBLEM STATEMENT

To build the backpropagation algorithm and tune hyperparameters of the neural network to optimize the validation accuracy.

II. NETWORK STRUCTURE

The hyperparameters which can be tuned are the following:

- Number of Layers
- Number of Neurons in each layer
- Learning Rate
- Batch Size
- Number of Epochs

Following are the comparisons between default network and tuned network:

HyperParameter	Default Network	Tuned Neural Network
No of Layers	1	1
Neurons in Layer 1	20	32
Learning Rate	0.001	0.002
Batch Size	128	64
Number of Epochs	100	75
Training Accuracy	90.02	94.46
Validation Accuracy	88.35	90.83
Training Loss	0.68232	0.55579
Validation Loss	0.79787	0.69680653
Test Accuracy	87.78	90.78

III. Derivative of Cross Entropy

The cross entropy function is given by:

$$L(\hat{y}, y) = -\sum y_c . \ln(\hat{y})$$

$$L(\hat{y}, y) = -(1-y).\ln(1-\hat{y}) - y.\ln(\hat{y})$$

We know that $h^{(k)} = y = \sigma(a^{(k)})$

We have to find the gradient to the loss:

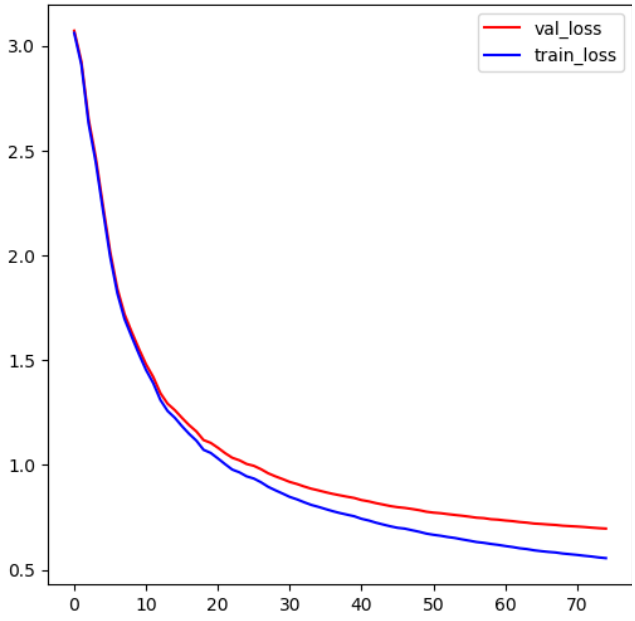
$$\frac{\partial L(\hat{y}, y)}{\partial a_j} = -(1-y_j)\ln\left(1-\sigma\left(a^{(k)}{}_j\right)\right) - y_j \ln\left(\sigma\left(a^{(k)}{}_j\right)\right)$$

$$\frac{\partial L(\hat{y}, y)}{\partial a_j} = y_j - \hat{y}$$

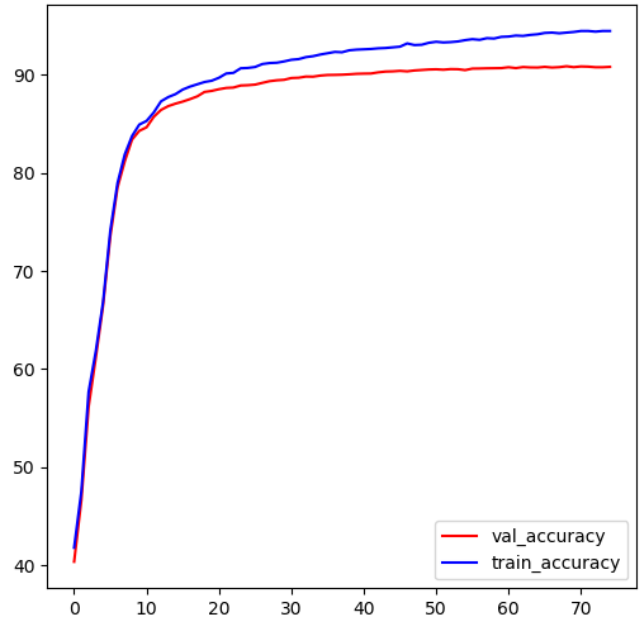
IV. Learning Curves

I have recorded training/validation accuracy and training /validation loss over number of epochs. Training and Validation curves have been recorded on the same graph for comparison of model performance between training set and validation set.

- **Loss Curve :**



- **Accuracy Curve :**



The number of epochs have been reduced from 100 to 75 as the accuaracy curve plateaus after 50 epochs and may lead to overfitting. The batch update size is also reduced from 128 samples to 64 samples to make frequent updates. The learning rate has been doubled in order to converge faster. The experiment data shows that learning rates greater than 0.005 leads to underfitting. Number of Neurons in layer 1 has been increased from 20 to 32 to accommodate more weights and increase model complexity

.

V. Test Set Accuracy

The tuned model gave approximately 4.46% better accuracy on training dataset, 2.5 % better accuracy on validation dataset and 3% better accuarcy on the test dataset. The accuracy on the test dataset is 90.7