Assignment 2nd

Soln (1) :- We have

$$y^k = X * W^k$$

$$y^k(p,q) = \sum_{i=0}^{S-1} \sum_{j=0}^{T-1} X(p+i, q+j)\, W^k(i,j) \quad \text{— (I)}$$

Now,

$$\frac{\partial J}{\partial W^k} = \frac{\partial J}{\partial y^k} \cdot \frac{\partial y^k}{\partial W^k} \quad \text{(As per chain rule)} \quad \text{— (II)}$$

Now, $$\frac{\partial y^k}{\partial W^k} = \frac{\partial}{\partial W^k}\left[\sum_{i=0}^{S-1} \sum_{j=0}^{T-1} X(p+i, q+j) W^k(i,j) \right]$$

$$\Rightarrow \frac{\partial y^k}{\partial W^k} = \sum_{i=0}^{S-1} \sum_{j=0}^{T-1} X(p+i, q+j) \frac{\partial}{\partial W^k}\left(W^k(i,j)\right)$$

$$= \sum_{i=0}^{S-1} \sum_{j=0}^{T-1} X(p+i, q+j) \cdot 1$$

So,

$$\frac{\partial y^k}{\partial W^k} = \sum_{i=0}^{S-1} \sum_{j=0}^{T-1} X(p+i, q+j) \quad \text{— (III)}$$

So, as per eqn (ii)nd and (iii)rd

$$\frac{\partial J}{\partial w^k} = \frac{\partial J}{\partial y^k} \cdot \sum_{i=0}^{S-1} \sum_{j=0}^{T-1} x(p+i, q+j)$$

which can also be written as.

$$\frac{\partial J}{\partial w^k} = \sum_{i=0}^{S-1} \sum_{j=0}^{T-1} x[p+i, q+j) \cdot \frac{\partial J}{\partial y^k} \quad \text{(iv)}$$

As per eqn (i)st the eqn (iv)th can be written as

$$\boxed{\frac{\partial J}{\partial w^k} = x * \frac{\partial J}{\partial y^k}}$$

Hence proved.

Comments on DNN for MNIST data set :-

(1) Increasing the learning rates improves the accuracy on the test data set and on training data set as well.

(2) Increasing Batch Size decreases the accuracy on the test data set and on training data set as well.

(3) Increasing the neurons on the hidden layer decreases the accuracy.

(4) ~~Training~~

Comments on CNN for MNIST data set :-

(1) Increasing the learning rate improves the accuracy on both training and test data set, however process gets slow with lower learning rate.

(2) Increasing the batch size improves the accuracy on both training and test data set.

(3) plots are ~~future~~ included in the project file.

(4) Training accuracy is around ≃ 97%.

__/__/____

Comments on CNN for CIFAR:-

(1) Training accuracy is around 80% as I have only used basic numpy etc. written the code without using Keras or pytorch.

(2) Increasing the batch size decreases the accuracy on both ~~train~~ training and test data set. plots are included in project file.

(3) Increasing the learning rates decreases the accuracy on both ~~train~~ training and test data set.