Experiment No: 02

Aim: To Implement Feedforward neural networks with Keras and TensorFlow.

Problem Statement:

a. Import the necessary packages

b. Load the training and testing data (MNIST/CIFAR10)

c. Define the network architecture using Keras

d. Train the model using SGD

e. Evaluate the network

f. Plot the training loss and accuracy.

Objectives:

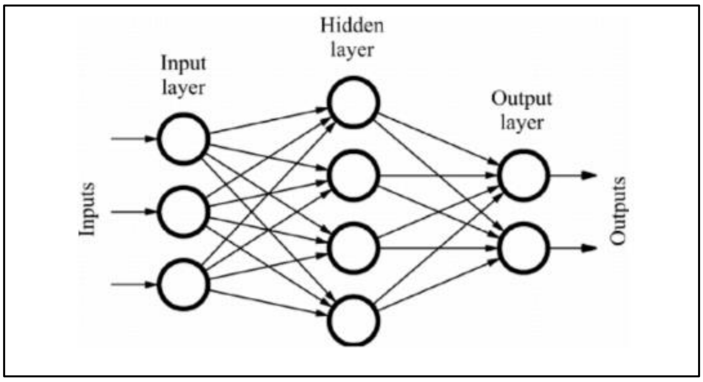
1) Apply feedforward neural network on MNIST/CIFAR dataset for classification

2) Evaluate the performance of models and plot the loss and accuracy graph

Theory:

A feedforward neural network is a type of artificial neural network in which nodes' connections do not form a loop. Often referred to as a multi-layered network of neurons, feedforward neural networks are so named because all information flows in a forward manner only. The data enters the input nodes, travels through the hidden layers, and eventually exits the output nodes. The network is devoid of links that would allow the information exiting the output node to be sent back into the network.

<u>The following are the components of a feedforward neural network:</u>



Layer of input

It contains the neurons that receive input. The data is subsequently passed on to the next tier. The input layer's total number of neurons is equal to the number of variables in the dataset.

Hidden layer

This is the intermediate layer, which is concealed between the input and output layers. This layer has a large number of neurons that perform alterations on the inputs. They then communicate with the output layer.

Output layer

It is the last layer and is depending on the model's construction. Additionally, the output layer is the expected feature, as you are aware of the desired outcome.

Neurons weights

Weights are used to describe the strength of a connection between neurons. The range of a weight's value is from 0 to 1.

Cost Function in Feedforward Neural Network

The cost function is an important factor of a feedforward neural network. Generally, minor adjustments to weights and biases have little effect on the categorized data points. Thus, to determine a method for improving performance by making minor adjustments to weights and biases using a smooth cost function.

The mean square error cost function is defined as follows:

$$\boxed{C(w, b) \equiv \frac{1}{2n} \sum_x \|y(x) - a\|^2.}$$

Where,

w = weights collected in the network

b = biases

n = number of training inputs

a = output vectors

x = input

‖v‖ = usual length of vector v

Loss Function in Feedforward Neural Network

A neural network's loss function is used to identify if the learning process needs to be adjusted.

The cross-entropy loss for binary classification is as follows.

Cross Entropy Loss:

$$L(\Theta) = \begin{cases} -log(\hat{y}) & \text{if } y = 1 \\ -log(1-\hat{y}) & \text{if } y = 0 \end{cases}$$

The cross-entropy loss associated with multi-class categorization is as follows:

Cross Entropy Loss:

$$L(\Theta) = -\sum_{i=1}^{k} y_i \log(\hat{y}_i)$$

Gradient Learning Algorithm

Gradient Descent Algorithm repeatedly calculates the next point using gradient at the current location, then scales it (by a learning rate) and subtracts achieved value from the current position (makes a step) (makes a step). It subtracts the value since we want to decrease the function (to increase it would be adding) (to maximize it would be adding). This procedure may be written as:

There's a crucial parameter η which adjusts the gradient and hence affects the step size. In machine learning, it is termed learning rate and has a substantial effect on performance.

a. The smaller the learning rate the longer GD converges or may approach maximum iteration before finding the optimal point
b. If the learning rate is too great the algorithm may not converge to the ideal point (jump around) or perhaps diverge altogether.

Conclusion:

Thus, we have implemented the Image classification using feed forward neural network model. During the execution of each iteration (epochs), our model's training and validation data accuracy increases and loss decreases however significant change in the accuracy is observed after 80th iteration (epoch).