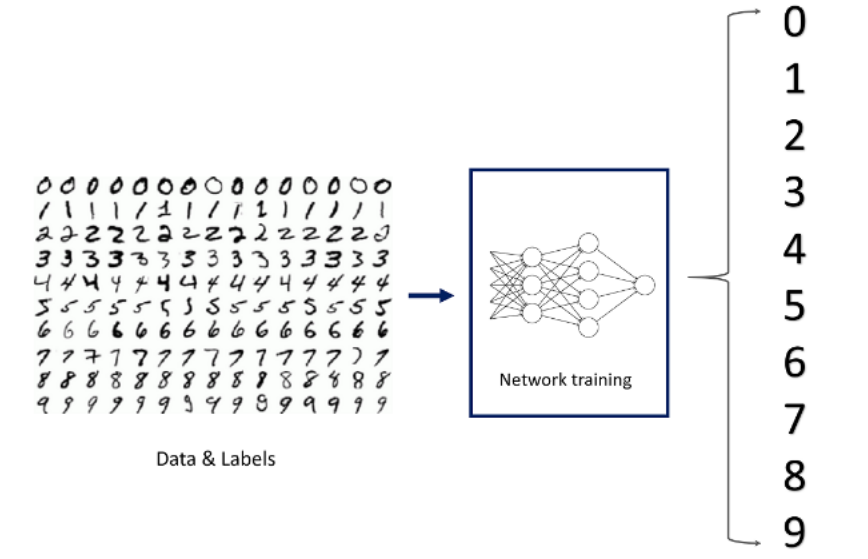
Summarizing MNIST Dataset

The most popular dataset used in neural networks is the MNIST (Modified National Institute of Standards and Technology) dataset. It contains 60,000 handwritten images of the digits from 0 to 9 in the training dataset. Each image is 28\*28 pixels of 0s and 1s (Pixels are scaled), where 0 stands for black and 1 stands for white. A model can be trained on this data to recognize hand-written numbers

A 28\*28 image gives 784 inputs, and the target variable has a value from 0 to 9, so the output layer will have 10 nodes, each node corresponding to a digit from 0 to 9. The number of neurons and hidden layers are the hyperparameters in Neural Networks that can be changed.



* The output from the above model will give 10 numbers, and each number represents the probability of the number being 0, 1, 2 … 9, and it is called predicted y.
* The predicted values should be close to the actual values, so to achieve this derivatives/gradients of predicted y with respect to all the weights are calculated.
* From derivatives, how predicted y changes or in which direction it changes with respect to each of the weights is determined.
* So, derivatives/gradients are used to adjust the weights so that the predicted y becomes close to the actual y.

Multiple perceptrons, when put together in a well-designed structure and more importantly trained with the relevant data in the right manner, can show signs of being intelligent.

The structure that is built above is known as Feed Forward, and it essentially feeds the input, and everything computed in a given layer is moved to the next layer, and so on. When every node in each layer is connected to every node in the next immediate layer, the structure is known as a Fully Connected Neural Network.

Happy Learning!!