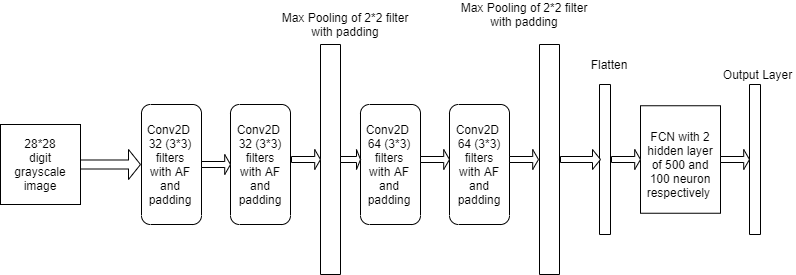
**CNN Network for MNIST Dataset with Different Activation Function (RELU, TANH and ELU)**

**Technical Documentation**

CNN is Convolutional Neural Network created on MNIST dataset (dataset with 70000 hand written digits 28\*28 grayscale images) tries to learn, identify and classify the handwritten digits using several filters and fully connected network. This method uses filters to identify the spatial relationship of each image pixel with respect to its surrounding pixels. It tries to learn the patterns in the images.

Technically speaking, the network uses back propagation method to learn different weights and biases for each layer and tries to optimize them. Back\_propagation function initializes the network by assigning random weights to the connections of input, hidden and output layers. Below is the diagram of CNN network used.



The CNN model defined has been trained for 50 epochs with batch size of 200.

Training Data Size : 42000

Test Data Size: 10000

Validation Data Size: 18000

The model built uses training dataset to train and learn with each of the activation function defined. Then model is evaluated with validation dataset. Each model history which contains the loss and accuracy for training and validation dataset is stored in history list. Each model score is then evaluated on test dataset and is stored in scores list. Then the model is used to predict the classes on test dataset and predicted values of each model is stored in predictions list.

Then a graph is plotted for training and validation accuracy and loss for each model with specific AF.

Then accuracy of each model is calculated from score for test dataset and is printed.

Then F1 score was calculated for each model with specific AF.

A graph is also plotted for accuracy and loss function vs epochs for training and validation dataset for each model.

All the metrics and graph has been defined in the model creation python file.

The metrics helps us understand the importance of choosing the right AF for a model.

From the model created, it is evident that the model with AF: “tanh” has the highest accuracy.