**Implementation Details**

* Dataset : Mnist Digit Classification

# build the network

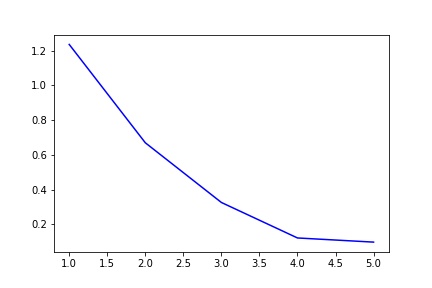
# w1/b1 w2/b2 w3/b3

#784(inputs) ---> 32 ---> 32 ---> 10(output)

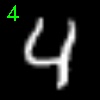
# x z1 a1 z2 a2 z3 a3=y

* Weights and Biases are initialized using random library(in between 0 to 1)
* Activation Function Used is Relu and Softmax at the output.
* As it is classification algorithm, Loss function used is categorical cross entropy
* Updated weights are stored in a pickle file and attached in the folder.
* I had kept the Batch size as 1 following stochastic batch gradient descent but the algorithm had return in a way that can take different batch sizes(Batch size =(8,16,32..) following Mini-Batch gradient descent).
* Number of Epochs taken as 5 and Test accuracy is 82.9% and Train Loss is 0.12. If Increase the number of epochs, the test accuracy gradually increases.

**Plot of Loss Function vs Epochs**



**Test Results:**

****

**Failures:**

****