DATE

Ans-3) For this problem of MNIST classification I have used the Keras library.

I have used the tanh activation in the hidden layer & the sigmoid activation for the outer layer.

The output of the network is the predicted class of the test.

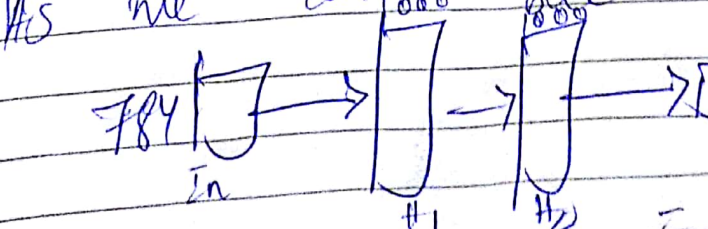
So with for this implementation I have optimised it with RMS prop & adam & observed that test accuracy is 100%.

This is due to fact that

(i) The classifier is non-linear (due to activation) so it easily classifies non linearly separable data.

(ii) The network is overparameterised.

As we can see



Parameters = 784 × 1000 + 1000 × 1000 + 1000 × 10
= 1785000 parameters

As this is a small dimension problem we are using overparameterised function & so we are getting accuracy is 100%.

PAGE