

For the problem, I have implemented a neural network using Sklearn library. I used 1000, 1000 layered 2 hidden layers with a relu activation function.

This was done as in the output layer, we want the value to binary 0 and 1.

The solver used is ~~logit~~ with a non-entropy loss and the network is extremely efficient at learning to classify the images. ~~also used 'adam'~~

- i) The classifier is non-linear so it works well even if data is not-linearly separable.
- ii) The model is overparametrized and likely overfitting the model, coz error in training data ~~100%~~ is 0% and also on test data is 0.5%.