Assignment 2

XOR Model

This model implements the nonlinear data of XOR gates. It has 1 hidden layer with sigmoid and tanh as the activation layer.

The model with the sigmoid activation function was easier to train than the model with tanh activation function. The model with tanh activation function converges to a solution while the model with tanh activation didn't.

Hand-written digit recognition

Implemented 4 different models with different hyperparameters.

1) Model with all zeros
   Here, model was trained with 1 hidden layer with 64 nodes. The learning rate was 0.001 and ran for 10 steps. The testing accuracy was 82.51%. The validation loss can be given as follow
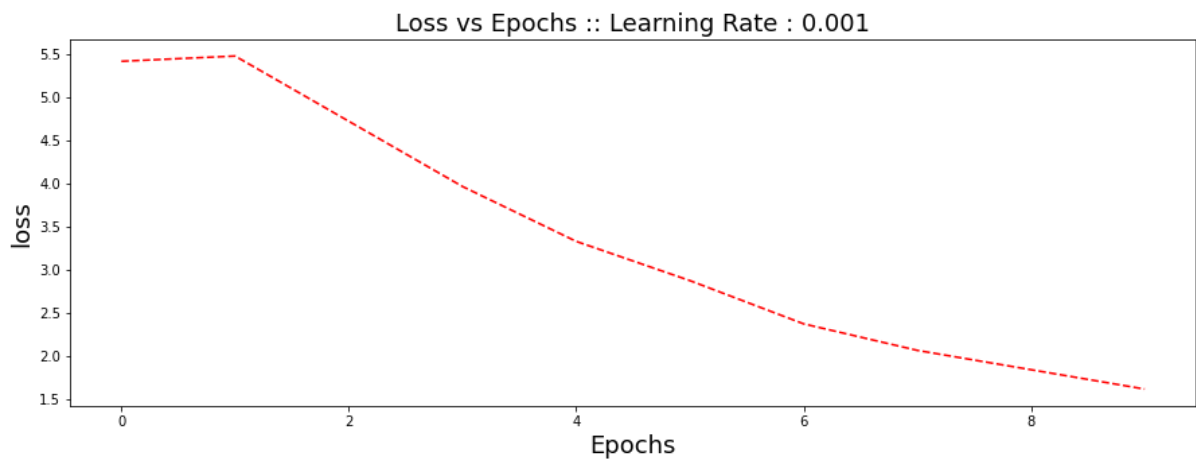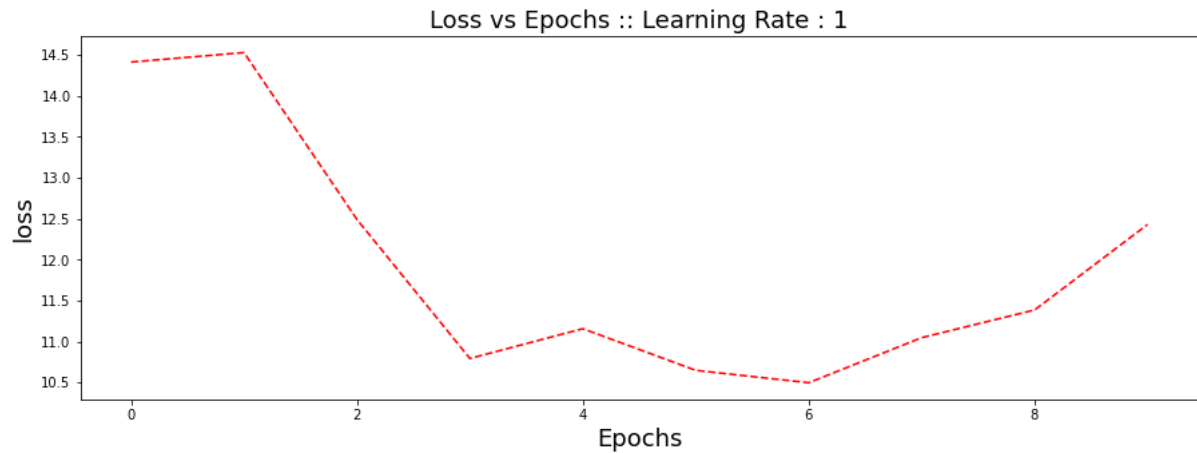

2) Models with weights between -10 and 10
   Here the model was trained with the weights ranging from -10 to 10. It contain 1 hidden layer with 100 nodes. The learning rate for the model was 0.01 and ran for 10 steps. The testing accuracy of the model was 89.21%. The validation loss can be given as follow
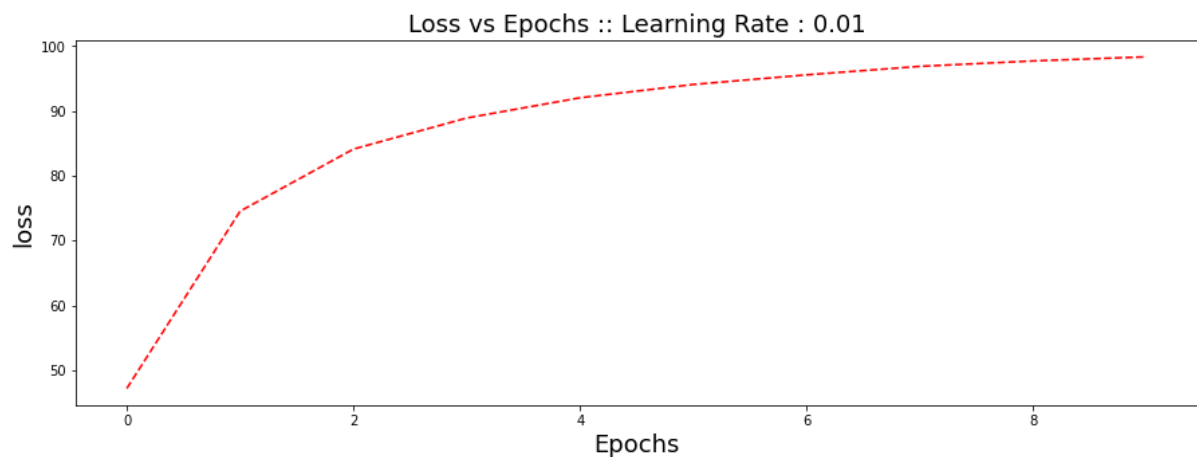
3) Models with learning rate 1

Here the model was trained with the weights ranging from -1 to 1. It contain 1 hidden layer with 100 nodes. The learning rate for the model was 1 and ran for 10 steps. The testing accuracy of the model was 84.86%. The validation loss can be given as follow

Loss vs Epochs :: Learning Rate : 1

4) Models with learning rate 0.001

Here the model was trained with the weights ranging from -1 to 1. It contain 1 hidden layer with 100 nodes. The learning rate for the model was 0.001 and ran for 10 steps. The testing accuracy of the model was 86.09%. The validation loss can be given as follow

Loss vs Epochs :: Learning Rate : 0.001