Can Dilsiz
21702688

GE 461 Spring 2023

Homework Assignment 3

ARTIFICIAL NEURAL NETWORK REGRESSOR

PART A:

+ **Is it sufficient to use a linear regressor or is it necessary to use an ANN with a single hidden layer? If it is latter, what will be the minimum number of hidden units?**

For this case it is not sufficient to utilize linear regressor due to the fact that provided data is more complex for linear regressor to represent. The scatter plot results provided in the Part B, will be enlighten the reason with visuals plots. ANN with a 4 hidden layer is adequate for predicting the output based on the given inputs. Back propagating algorithm helps to determine the weights by extracting output-input successors weights.

+ **What is a good value for the learning rate?**

The learning rate is a hyper parameter in the training process of an Artificial Neural Network (ANN) that decided the step size at which the model updates its parameters during gradient descent. It plays a crucial role in the convergence and stability of the training process for the implemented network. Choosing a good value for the learning rate is important to ensure effective training and provide good stability.

Good value for learning rate determined as 0.00001

+ **How to initialize the weights?**

Weight initialization in Artificial Neural Networks (ANNs) is a crucial step that can impact on the convergence, training speed, and avarage performance of the network. There are various weight initialization methods, my implementation includes random uniform weight that weights are determined randomly between provided interval [a,b] that both a, b < 1.

Between [-0.2 0.2] is good value for initial weights but, certain interval is given in **Part B.**

+ **How many epochs should you use? How to decide when to stop?**

The number of epochs to use in training an Artificial Neural Network (ANN) is a crucial hyper-parameter that determines how many times the entire dataset is passed forward propagation and backward propagation through the network constructed during the training. The number of Epochs is decided using trial and fail method and the appropriate number is decided as 80,000. The algorithm will stop entering the data set using forward and backward propagation, when the epochs end that is 80,000.

+ **Does normalization affect the learning process for this application?**

Normalization refers to the process of scaling and standardizing the input data to a common range, typically between 0 and 1 or with a mean of 0 and standard deviation of 1. Thus, for my implementation normalization helped for better stability and convergence. Without using normalization, worse results has been occurred.

Can Dilsiz
21702688

PART B:

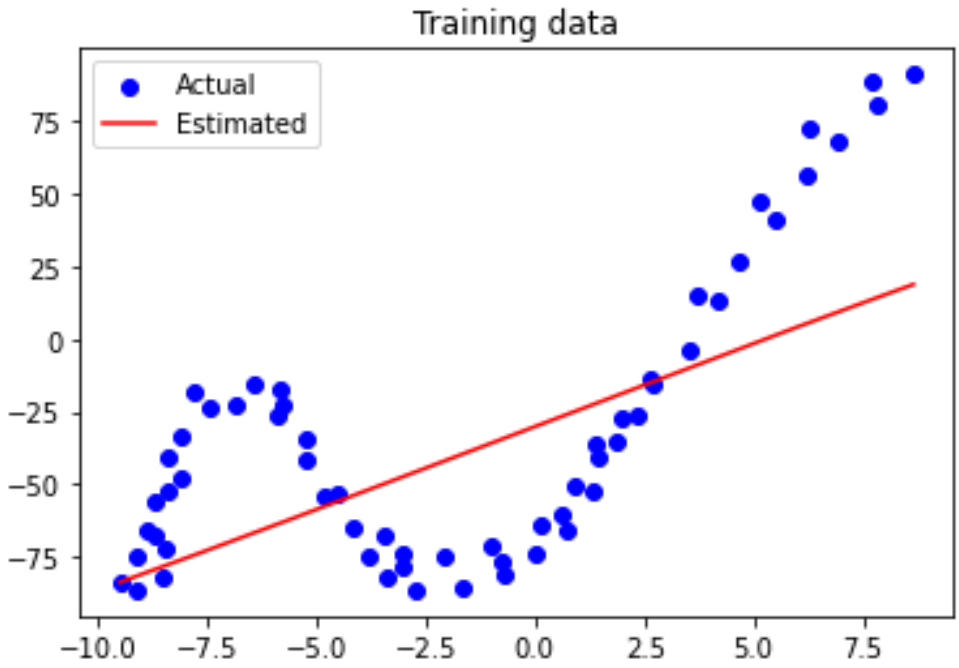

Figure 1.1: ANN Training Data with Prediction Curves

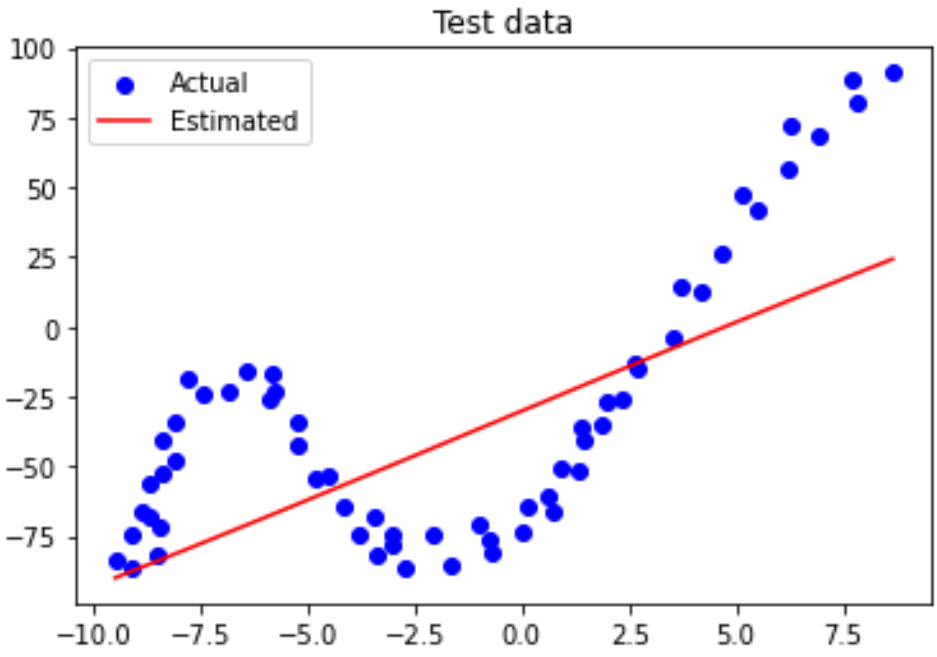

Figure 1.2: ANN Testing Data with Prediction Curves

+ **ANN Used:** 4 hidden units are used.
+ **Learning Rate:** 0.00001
+ **Range of Initial Weights:** Weights are distributed uniformly between [0.1 - 0.2]
+ **Number of Epochs:** 80,000
+ **When to Stop:** When Epochs End
+ **Normalization:** Yes —> (data - mean) / std
+ **Tranin Loss:** 0.47887574
+ **Test Loss:** 0.92083927

PART C:

Hidden Units	Learning Rate	Train Loss Avg.	Train Loss Std	Test Loss Avg	Test Loss Std
2	1e^05	0.6382	0.7672	0.9015	0.9766
4	1e^05	0.4278	0.5249	0.9821	0.8414
8	1e^05	0.4155	0.4012	1.0225	0.8349
16	1e^05	0.4134	0.3675	1.0345	0.8213
32	1e^05	0.4053	0.3561	1.0324	0.8259

Table 1.1: Training and test set losses.

Figure 1.3: Output Training Set with Hidden Units [2,4,8,16,32].