Assignment 3

COMP 4107– Assignment 3

Due: Thursday, 11 Mar 2021

Author: Group 29: Devon Robitaille, Alex Fillips

Carleton University

# Question 1

## Part (a) – Network Design

1. Reference supplied python code to see how *GradientDescentOptimizer* was implemented.

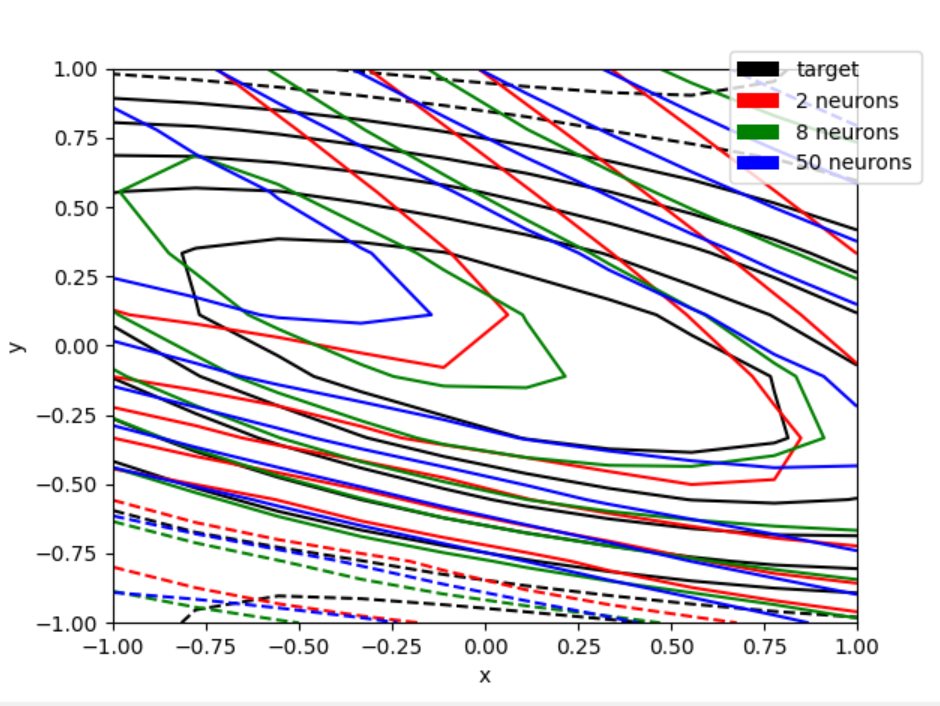


Figure - Function Contours

|  |  |  |
| --- | --- | --- |
| **# of Neurons** | **MSE / Loss** | **Epochs** |
| 2 | 0.08053534477949142 | 2289 |
| 8 | 0.07150878012180328 | 434 |
| 50 | 0.05533287301659584 | 932 |

1. For our experiments we used the “*tanh*” function as part of the TensorFlow library for the activation function of the hidden layers.

## Part (b) – Training

1. I set the total number of epochs to 2000 for this question to see at which point the optimizers would converge towards MSE of 0.02.

|  |  |
| --- | --- |
| **Model** | **Epochs to convergence** |
| Gradient Descent | 311 |
| Momentum | (Started at 0.001) Therefore, 0 |
| RMSProp | 40 |

1. MSE against epoch

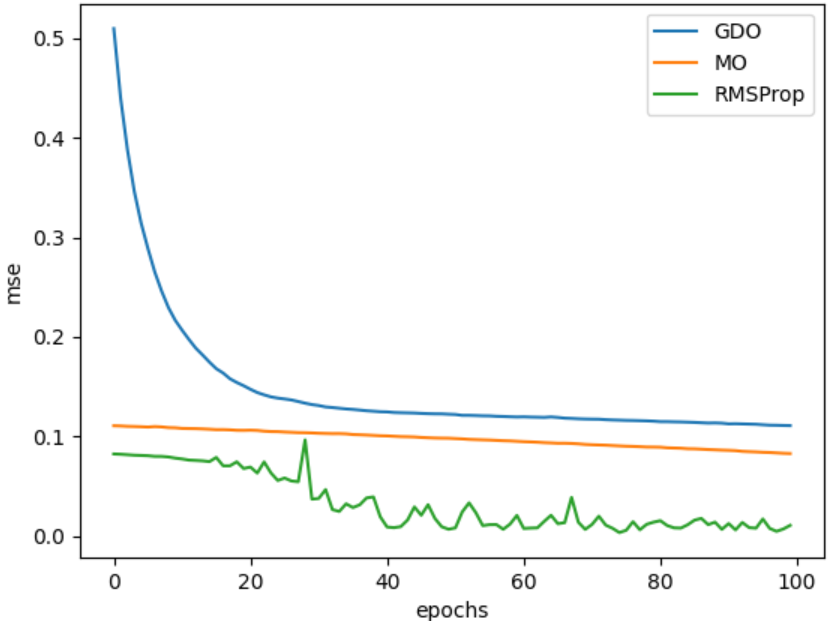


Figure - MSE against epoch number

1. Bar graph cpu time against 3 methods

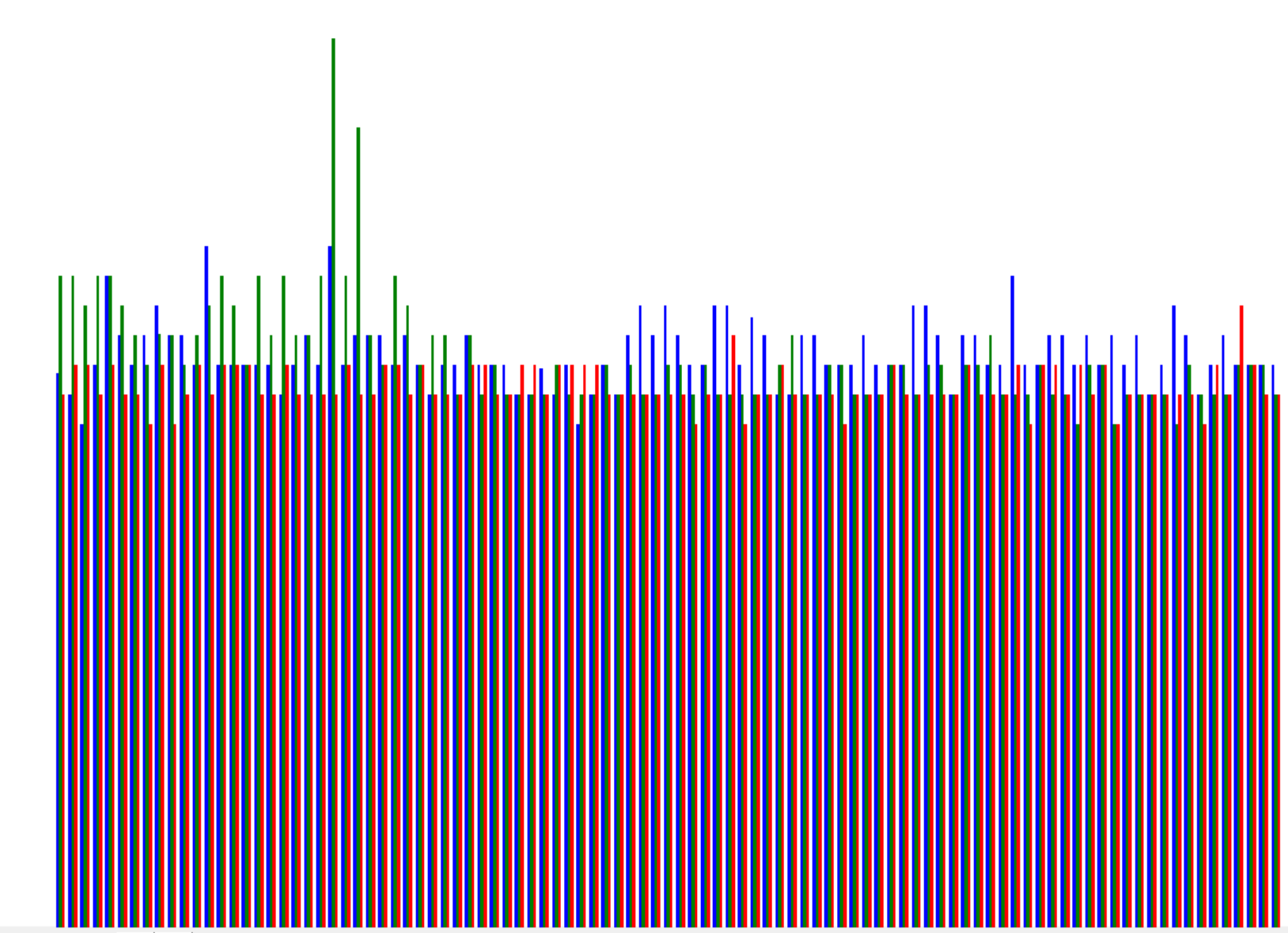


Figure - CPU time against 3 methods

1. Best accuracy at the end of 100 epochs

Based on the supplied *Fig.2* supplied in question 2, we can see that as the optimizers approached epoch 100, RMSProp provided the highest level of accuracy which can be visualized by the amount of loss or MSE.

1. Most accurate when the training error is reached

|  |  |
| --- | --- |
| **Model** | **Accuracy at end of training** |
| Gradient Descent | 0.06798079 |
| Momentum | 0.06798079 |
| RMSProp | 0.06798079 |

Since all three models provided the same level of accuracy by taking the average of the prediction against a range of possible values. It would say that they are all equally accurate.

## Part (c) – Early Stopping

1. Confirm that approximately 8 neurons are a good choice for the current problem

Using all of the previous data that has been calculated along with the interpretation taken from the different plots, we can say that using approximately 8 neurons appears to be the best choice for the current problem. By using 8 hidden neurons of the neural network, we can get the best mix of total number of epochs to train against the accuracy of the network which can be measured in MSE. Mainly that we were able to reach the desired MSE threshold of 0.02 most consistently which 8 neurons. Following this, we can see from the graph supplied under “run experiments” *Fig.4* that having 8 neurons seems to fit around average while also reaching this state a lot sooner than its counter parts.

1. Run experiments…

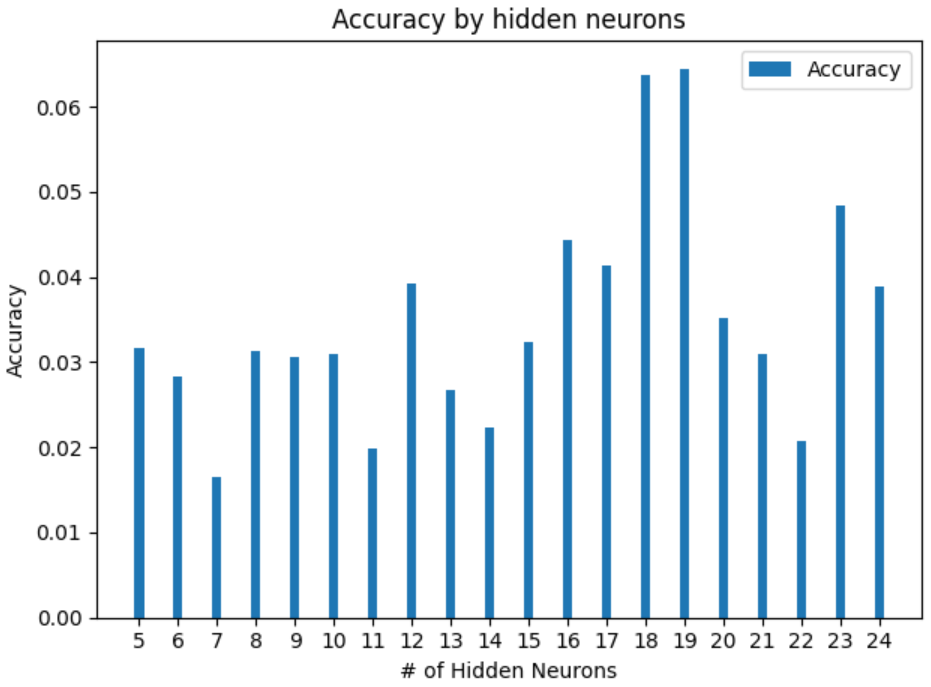


Figure - Evolution of Accuracy for 500 Epochs

1. Figure 6 and 7

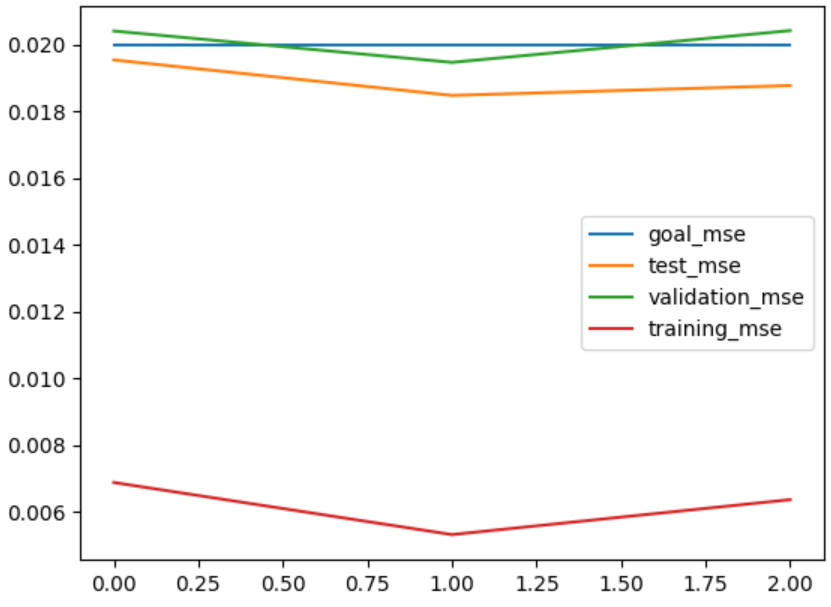


Figure - Evolution of the MSE

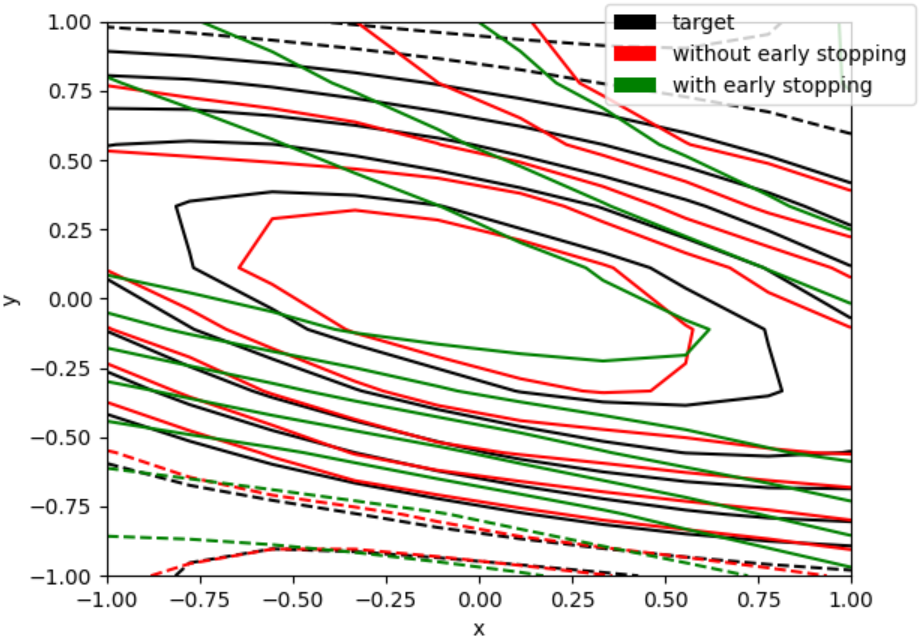


Figure - Function Contours