Gradient Descent and Backpropagation – HW3 Report

Agathiya Tharun, *Graduate Student*, MAE 263F, *Mechanical and Aerospace Engineering*

***Abstract* — This report details the results of a linear fit and nonlinear fit simulation to highlight the effects of gradient descent and backpropagation. A Mean Squared Error (MSE) is used to compute loss and hyperparameters are varied to identify their effects on the simulation.**

I. Hyperparameter Discussion

The optimized model that minimizes loss was achieved by tuning hyperparameters, namely the learning rate and number of epochs.

It is desired to find a balance for these values. The learning rate determines how quickly a model learns and how it adjusts its weights during training in response to the gradient of the loss function. It is analagous to step size and thus requires a balance. Too high a learning rate can lead to inaccuracies in learning, overshooting, or failure to converge. High learning rates also demand more processing power. In the contrary, too low of a learning rate would take the model longer to converge and learn. However, convergence would be more precise, but this would be at the expense of spending more resources at suboptimal steps.

At each epoch, the entire training data set is reviewed. Each epoch represents an iteration of learning. Too few epochs could lead to a lack of learning but quicker convergence, although inaccurate. The model would experience large loss and underfitting. Too many epochs could lead to lower loss but only for the training data – the model may experience overfitting.

In the context of gradient descent, these hyperparameters affect the tuning ability of the model. The neural network identifies the gradient of the loss with respect to its weights and adjusts them accordingly through backpropagation at each epoch in the opposite direction as to minimize the loss.

II. Assignment 1

A random seed is used to generate linear data points with some Gaussian noise. It is desired that a linear fit model with gradient descent is used as part of a backpropagation-based loss function to minimize error and best represent the generated values with a line of best fit.

1. Data Comparison

The comparison between the sample data and the line of best fit created from the predicted data can be seen below in *Figure II.1.1*.

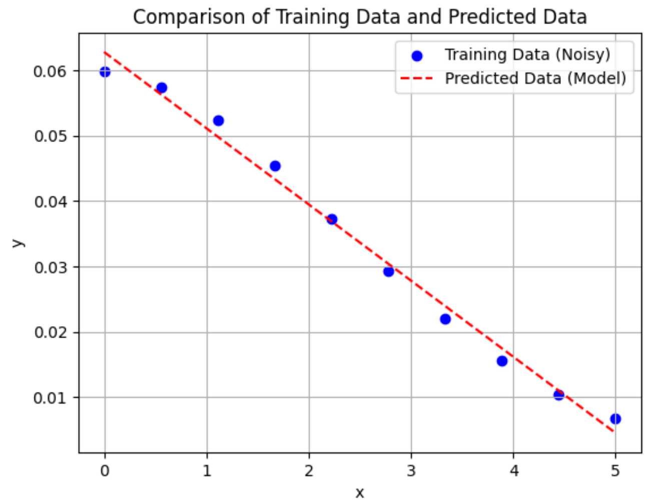

Figure II.1.1: Sample vs. predicted data comparison for a linear fit

2. Hyperparameter Tuning

The optimized model was produced at a learning rate of 0.001 and 10000 epochs. Higher or lower values led to a lack of convergence or increased loss. This hyperparameter configuration led to a minimum loss value of $3.48 * 10^{-6}$.

An example of this can be seen when a learning rate of 0.00001 and 100000 epochs are used. The minimum loss value for this model was 0.000348688. The results of this model's accuracy can be visualized below in *Figure II.2.1* below.

Figure II.2.1: Unbalanced hyperparameters can lead to inaccurate models

As seen above, poor tuning of hyperparameters can lead to underfitting as the gradient loss function isn't utilized properly; this affects its convergence.

III. Assignment 2

A random seed is now used to generate nonlinear data points with Gaussian noise. It is desired that a nonlinear fit model with gradient descent is used as part of a backpropagation-based loss function to minimize error and best represent the generated values with a best-fit curve.

1. Data Comparison

The comparison between the sample data and the best-fit curve created from the predicted data can be seen below in *Figure III.1.1*.

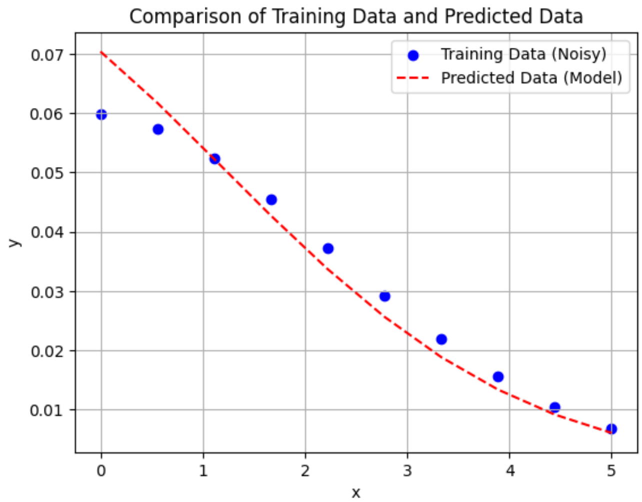

Figure III.1.1: Sample vs. predicted data comparison for a nonlinear fit

2. Hyperparameter Tuning

The optimized model was produced at a learning rate of 0.001 and 10000 epochs. Higher or lower values led to a lack of convergence or increased loss. This hyperparameter configuration led to a minimum loss value of 00016592.

An example of this can be seen when a learning rate of 0.0001 and 100000 epochs are used. The minimum loss value for this model was 0.00248731. The results of this model's accuracy can be visualized below in *Figure III.2.1* below.

Figure III.2.1: Unbalanced hyperparameters can lead to inaccurate models

As seen above, poor tuning of hyperparameters can lead to underfitting as the gradient loss function isn't utilized properly; this affects the gradient descent convergence of the nonlinear model.

References

[1] Professor M. Khalid Jawed, MAE 263F https://bruinlearn.ucla.edu/courses/193842/modules