Homework 2 Report

Arrian Chi

Abstract

In this homework, I reviewed the concepts of linear and nonlinear regression and applied them to the data set.

PROBLEM 1

a:

Fig. 1. Plot of predicted values of y_{pred} versus actual values of y_{actual} (linear)

b:

The number of epochs is the number of times the gradient descent will be applied, while the learning rate controls how much the prediction changes at each step. Increasing the number of epochs will slow down the performance of the model (takes more time to iterate) while increasing the learning rate will make the model step faster. However if the learning rate is too high, the model may become unstable and not converge. If the learning rate is too low (relative to epochs), the model may take a long time to converge and still not have converged at the end of the operation. If the epochs is too high, it is possible that the model will overfit the data or do redundant fitting, but if the epochs is too low, the model may not have enough time to converge.

The best combination that minimizes loss the quickest with a final loss $= 0.0000034267728104$ are epochs $= 1000$ and learning_rate $= 0.1$. Even for smaller learning rates and larger numbers of epochs, the model converges to this final loss (it can't get any better).
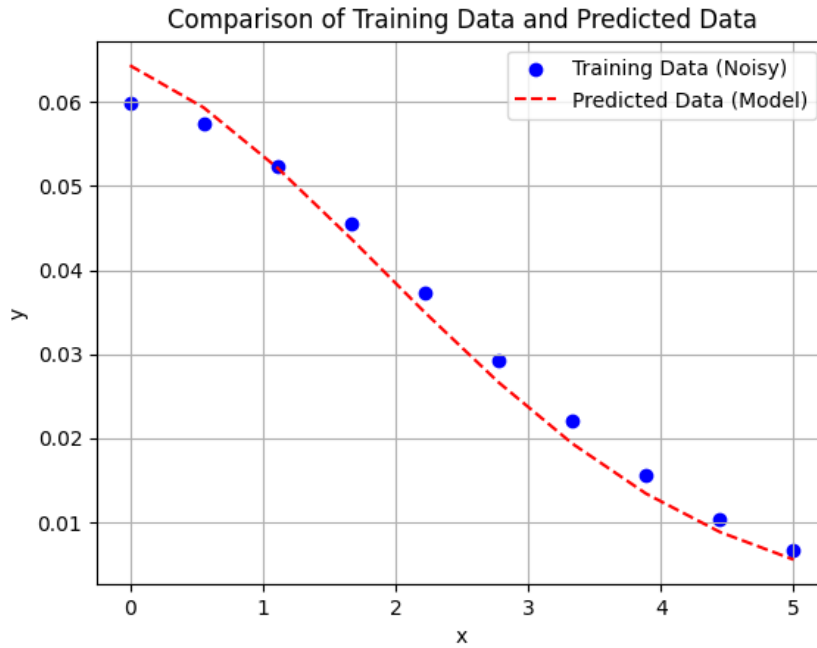
a:

Fig. 2. Plot of predicted values of y_{pred} versus actual values of y_{actual} (nonlinear)

b:

The previous section about epochs and learning rates is applicable, however I found that if the epochs or the learning rate is too high, the model will start diverging (we see the loss increasing/we start seeing NaNs!). I also found that a good initial guess for the parameters is important for the model to converge because it is possible for the model to converge onto local minima. Algorithms like simulated annealing could help minimize the chance of this happening (finding the global minimum).

The best combination that minimizes loss with a final loss = 0.000005445874 are epochs = 96200 and learning_rate = 0.0001. Beyond this, the model starts diverging (stepping away from this minima).