Artificial Neural Network

COMP 60012

Janhavi Revashetti, Ivan Severinov

Introduction to Machine Learning

Imperial College
London

November 22, 2024

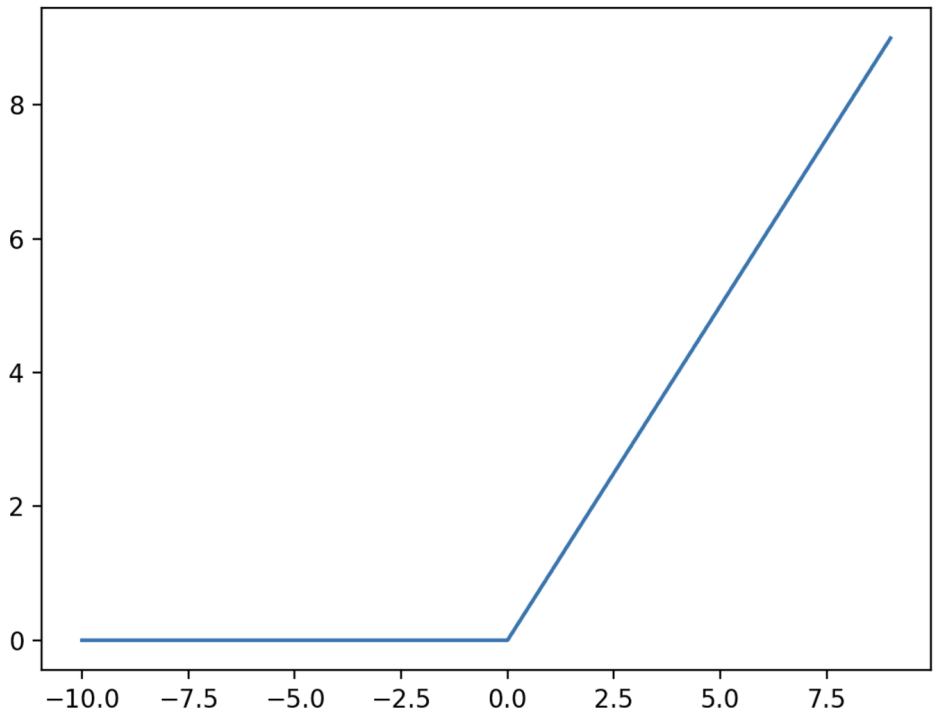

Figure 1: ReLu Activation Function

Description of the Model and Justification

For this assignment, we analyzed the California House Prices Dataset from the 1990 Census and implemented a neural network architecture to predict median house value based on certain attributes such as location, income, value, and many more. This model identifies learning patterns and optimizes through training and predicts the median house value. The ReLu model is a good choice because it is efficient and allows modeling for complex relationships in our data like ocean proximity and median income.

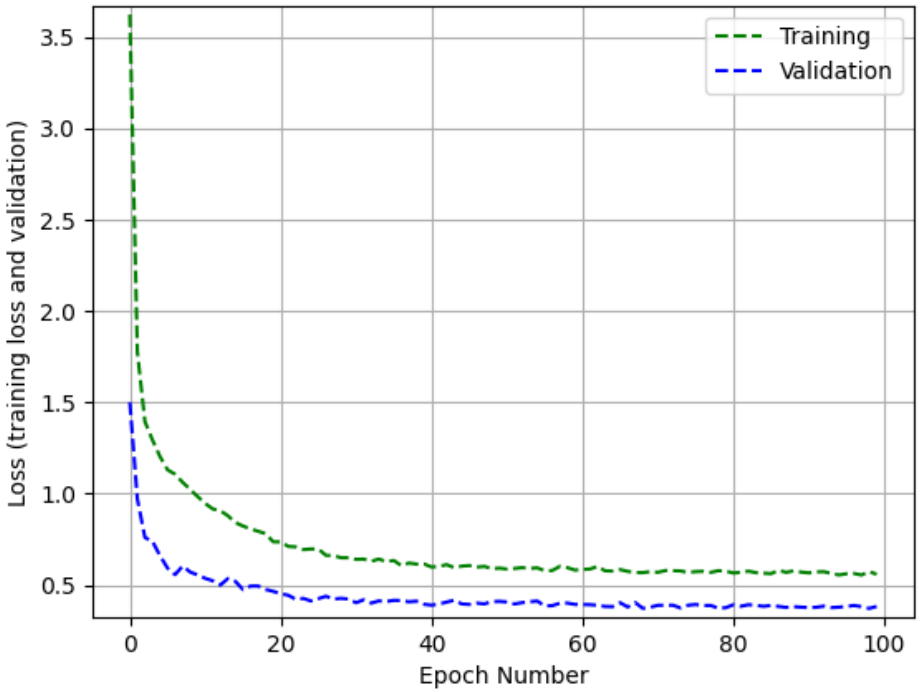

Figure 2: Training Loss vs Validation Loss for Every Epoch

Description of the Evaluation Setup

Two statistical methods were used to evaluate our model's performance: mean squared error and the R^2 Coefficient. The mean squared error squares the difference between the estimated and true values which therefore accounts for larger errors with a greater emphasis. A lower mean squared error demonstrates that the model's predictions are far more accurate than a model with a higher mean squared error.The Coefficient of Determination R^2 was used to determine how well the model also made accurate predictions but rather by determining the variance in the target variable. Typically a value closer to 1 means that the model predicted values with a high accuracy.

Model	Epoch	Learning Rate	Batch Size	Hidden Layers	MSE
Best Given Model	100	0.0005	32	16	0.3133518695831299
Best Hyperparamters	50	0.01	32	8	3.344752550125122

Figure 3: Comparing MSE Values for a Given Model vs Optimal Hyperparameter Model

Information about the Hyperparameter and Findings

The Hyperparameter search was used to find the optimal learning rate, batch size, hidden layers, and number of epochs through a basic graph search. Through this we were able to iterate through every Hyperparameter combination in order to find those with the best MSE score. In figure 3, we found that a smaller number of epochs, a higher learning rate, same batch size, and less hidden layers gave us the optimal hyperparameters for a better model from the given inputs for each variable.

Final Evaluation

Our best given model was found using 100 Epochs, a significantly small learning rate, an average batch size, and a small amount of hidden layers. We found that we received an MSE of 0.31335 or approximately 31000 dollars on our Best Given Model. Since our MSE was significantly small in comparison to the target variable, our model's prediction is highly likely to be very accurate.