

# Homework 2 Report

In this homework assignment, we wanted to optimize performance of a neural net by trying various combinations of hyperparameters and evaluating their results.

We trained the model using three activation functions (logistic, tanh, and relu), two learning rates (0.01 and 0.1), and two configurations for the number of hidden layers (2 and 3). The table we provided (called 'Model\_Performance.csv') shows us the average of our performance log (called 'performance\_log.csv').

## Best Average in Model\_Performance.csv:

Activation: Relu • LR=0.01, • Epochs=200, • Layers=3	Train Accuracy: 0.923	Train Error: 0.077	Test Accuracy: 0.912	Test Error: 0.088
---	-----------------------	--------------------	----------------------	-------------------

Through the Model\_Performance and performance\_log files, we can find that the best activation function is relu. It consistently performed the best across all configurations. Relu performed as amazingly as it did because it uses non-linearity. This avoids the vanishing gradient problem and allows it to learn faster.

Logistic was pretty close to relu. It squashes input values into the range (0, 1). This is useful for binary classification but can lead to vanishing gradients.

Tanh was also pretty close. It squashes inputs between -1 and 1. It's zero centered, which can help in cases where data contains negative values. This solves the issues of logistic by outputting negative values but it unfortunately still suffers from vanishing gradient at extreme input values.