Notes about ea with NN: Increasing the population size from 100 to 1000 improved the results on the training data from 8 to 3, and from 63 to 44 on the evaluation. This is probably due to the fact it is able to find a reasonably good individual in the first generation. Increasing the number of hidden layers to (30,15,10,2) also incurred a massive performance cost, but did not increase the performance of the best individual by much. Also, the algorithm appears to converge after 20 generations, running it for longer only yielding a increase in the average fitness (tested up to 100 generations).

Notes about gp:

The symbolic regression performed better with a bigger initial population. I hypothesize this is because it had more genetic diversity to evolve. Increasing the sizes of the initial tree’s was to taxing to run, and only provided marginal improvements. Decreasing the parsimony pressure caused the tree’s to grow very large and extremely quickly, with only marginal improvements in the accuracy on the evaluation set. (I was able to get it to miss only 1 in the training set, but it roughly performed the same on evaluation). A larger parsimony pressure yielded much smaller trees, that were still close in accuracy (missed roughly 60). This could be useful in further research to figure out how the nodes relate to the data.

Comparison:

The GP and the Ea with NN(Referred as NN from now on) performed roughly the same in evaluation. With the GP only correctly guessing two more. They also had roughly the same amount of individuals. The GP was able to perform decently well with changes to it’s parameters, while the NN varied rather wildly. This suggests the GP is better suited for this sort of problem, which is probably due to it’s ability to model non-linear relationships well. Future work on this project could be comparing which cases they missed, and whether they intersected at all.

Comparison Data:

Total cases: 569 (212(M),357(B))

Training: 100 (50(M),50(B)

Ea with NN:

Missed: 40(evaluation) + 4(training)

Total percent correct: 92% aprox

GP:

Missed: 38(evaluation) + 2(training)

Total percent correct: 93% aprox