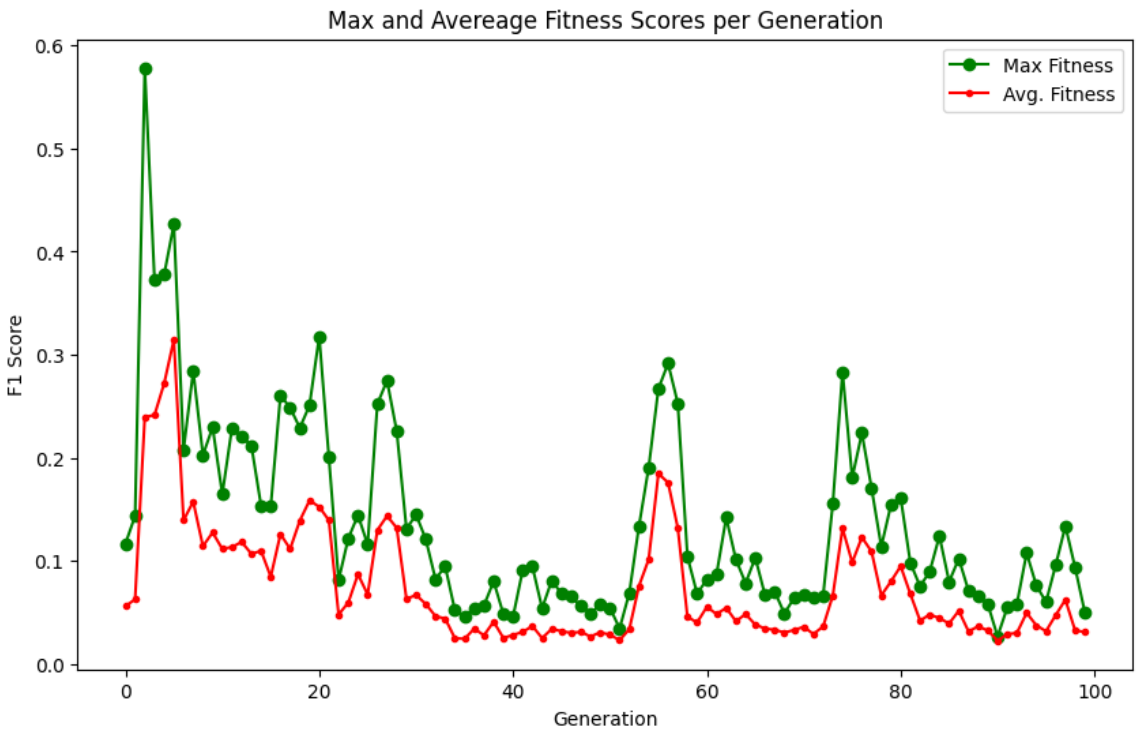
David McDevitt
MLDS Reinforcement Learning/Advanced Algorithms
HW4 Report
12/7/2023

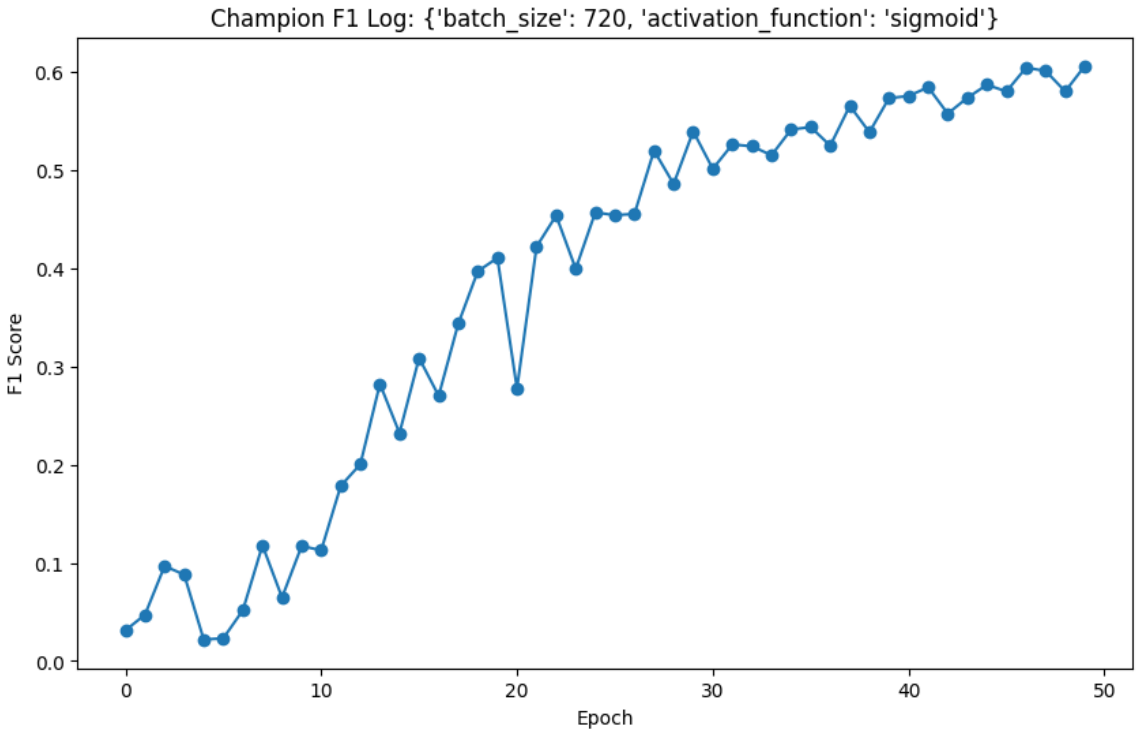
Part 1 - Genetic Algorithm

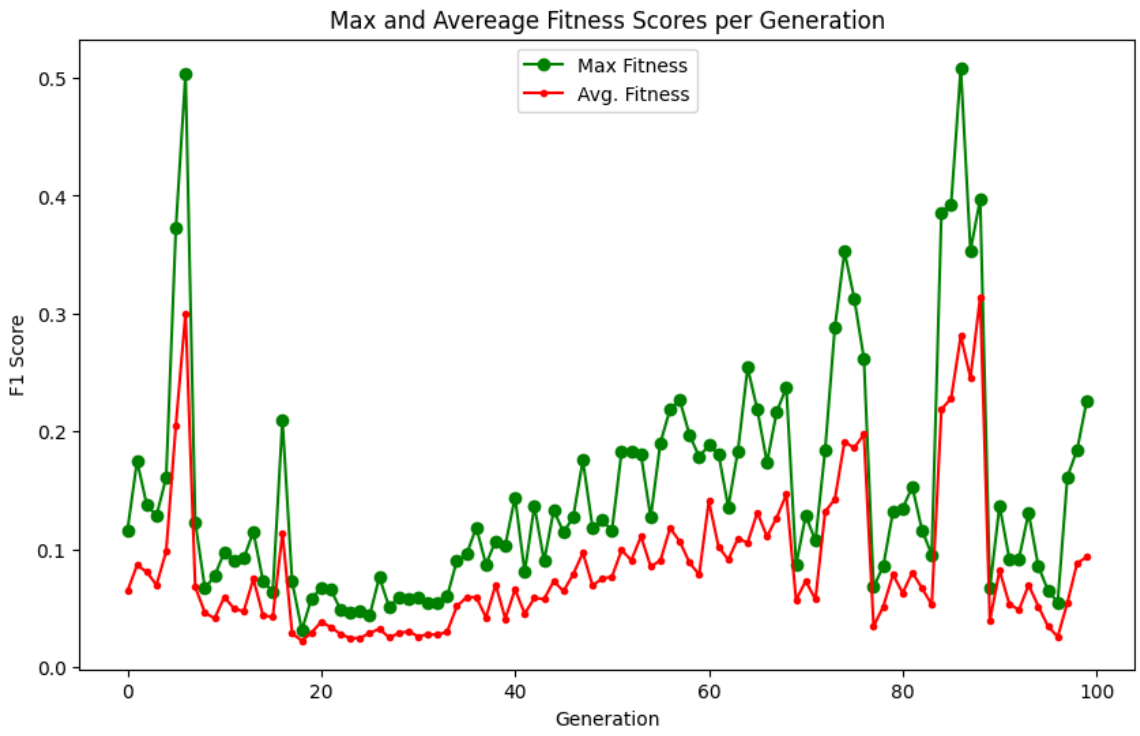
Population: 10
Number of parents per generation: 4
Number of generations: 100

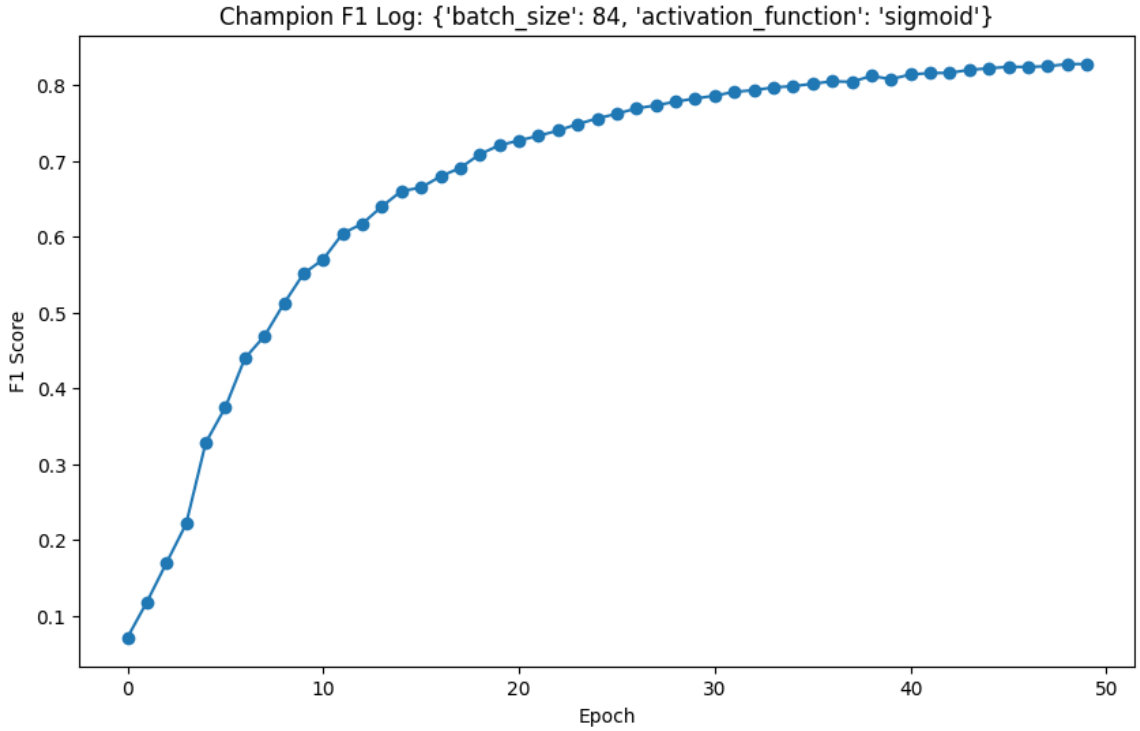
Fitness was evaluated with F1 scores taken after one epoch of training. The genome consists of variations of activation function (RELU, sigmoid, tanh) and batch size (integers ranging from 16 to 1024 images). Test F1 Score evaluated after 50 epochs of training from a newly created model under the same genome.

Run	Activation	Batch Size	Test F1
1	Sigmoid()	720	0.61
2	Sigmoid()	84	0.82
3	RELU()	713	0.72

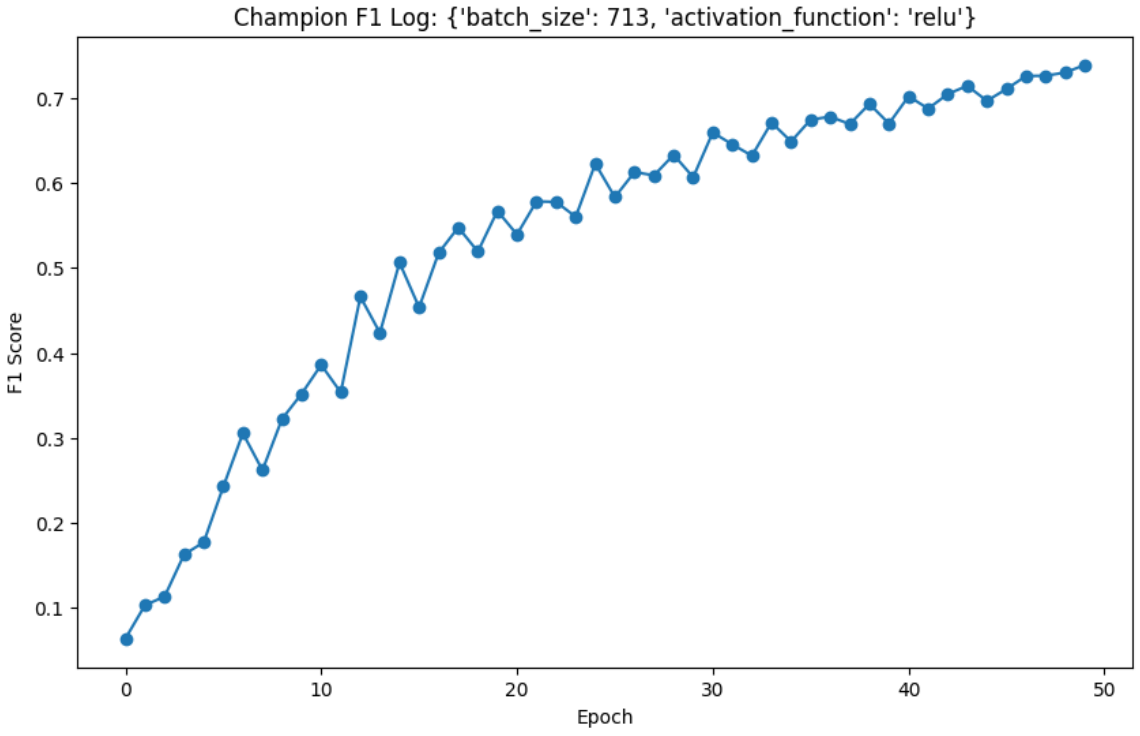

Max and Avereage Fitness Scores per Generation
Max Fitness
Avg. Fitness
F1 Score
Generation
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
20
40
60
80
100


Champion F1 Log: {'batch_size': 720, 'activation_function': 'sigmoid'}
F1 Score
Epoch
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
10
20
30
40
50


Max and Avereage Fitness Scores per Generation
Max Fitness
Avg. Fitness
F1 Score
Generation
0.0
0.1
0.2
0.3
0.4
0.5
0
20
40
60
80
100


Champion F1 Log: {'batch_size': 84, 'activation_function': 'sigmoid'}
F1 Score
Epoch
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0
10
20
30
40
50

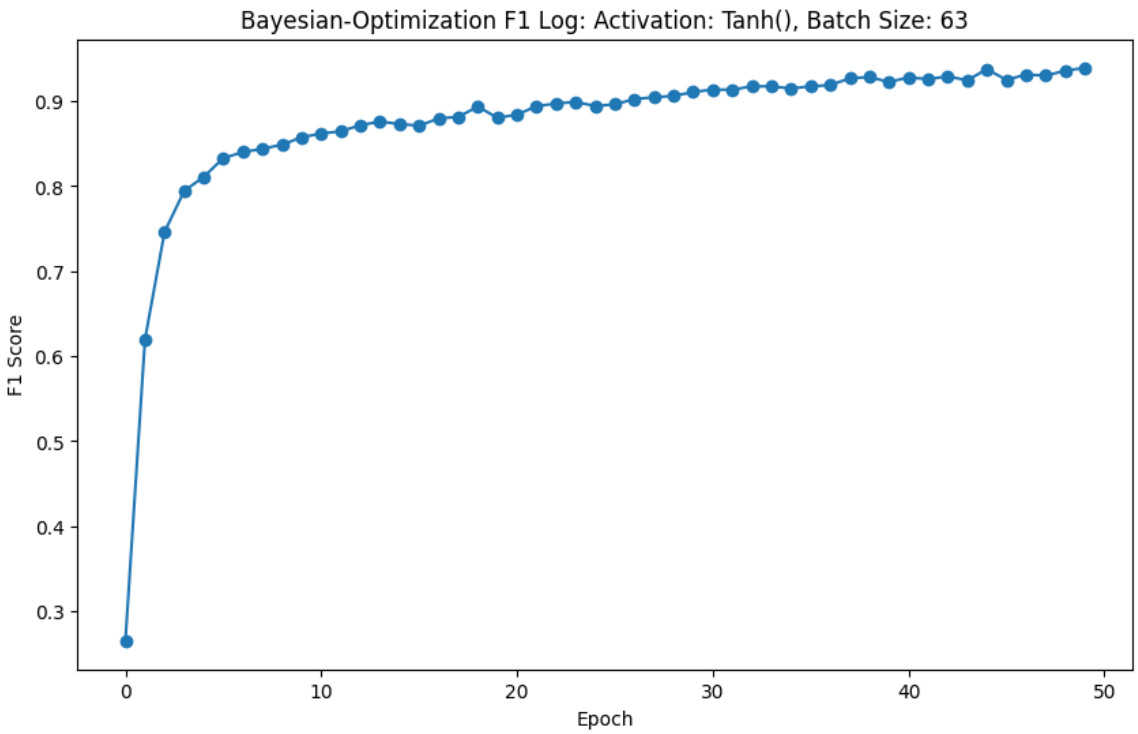

Max and Avereage Fitness Scores per Generation
Max Fitness
Avg. Fitness
F1 Score
Generation
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
20
40
60
80
100

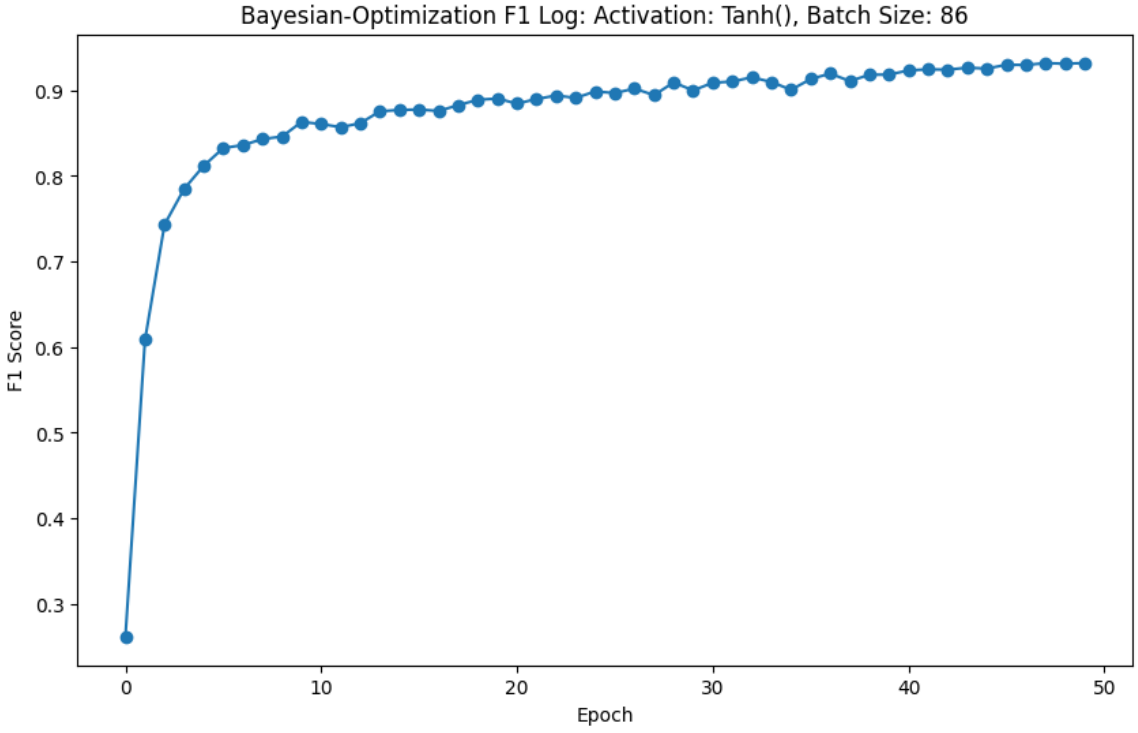

Champion F1 Log: {'batch_size': 713, 'activation_function': 'relu'}
F1 Score
Epoch
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
10
20
30
40
50

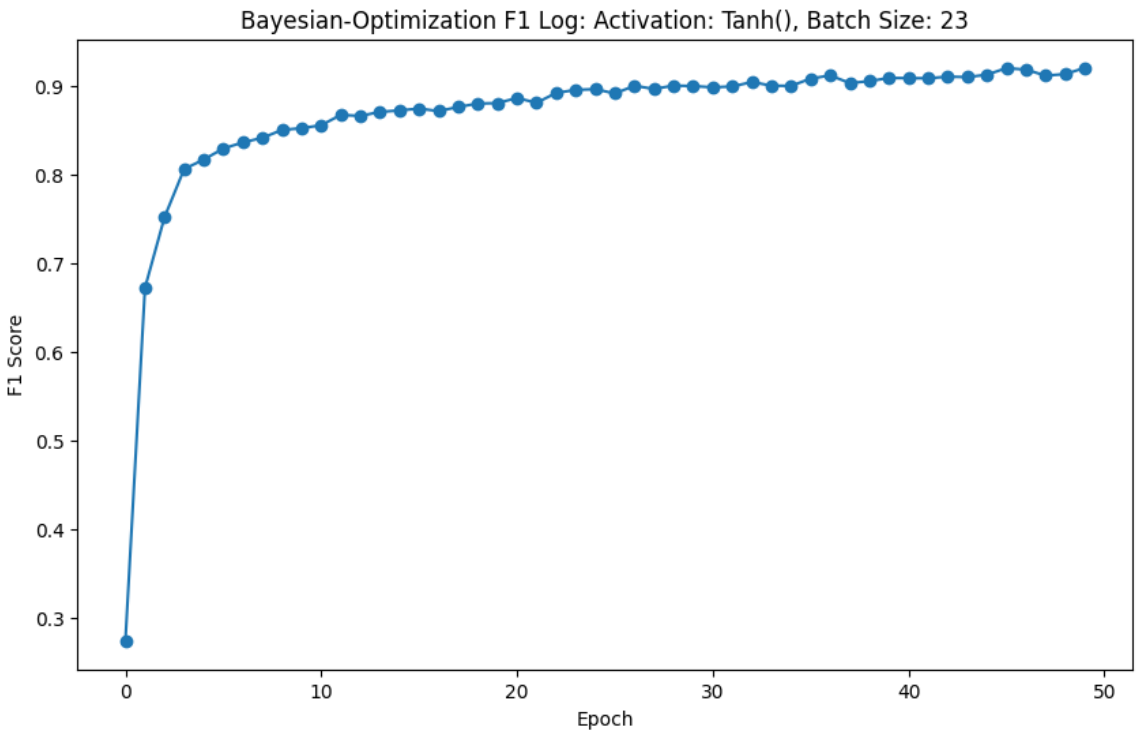
Part 2 - Bayesian Optimization

Iterations: 100

Run	Activation	Batch Size	Test F1
1	Tanh()	63	0.91
2	Tanh()	86	0.93
3	Tanh()	23	0.92


Bayesian-Optimization F1 Log: Activation: Tanh(), Batch Size: 63
F1 Score
Epoch
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
10
20
30
40
50


Bayesian-Optimization F1 Log: Activation: Tanh(), Batch Size: 86
F1 Score
Epoch
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
10
20
30
40
50



Comparison

Bayesian optimization chose a more consistent set of hyperparameters and was more efficient at reaching a solution than the genetic algorithm implementation. The small number of training iterations used to evaluate models in the population simulation was insufficient for the task of consistently choosing the optimal hyperparameters, and increasing the number of iterations would be impractical. Different parenting methods could also be explored to ensure that parents create children that are generally at their level of fitness, but that task would be domain specific.